MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 110 No. 4 1950

Published and Sold by the
ROYAL ASTRONOMICAL SOCIETY
BURLINGTON HOUSE
LONDON, W. 1

Price Nine Shillings

NOTICE TO AUTHORS

1. *Communications.*—Papers must be communicated to the Society by a Fellow. They should be accompanied by a summary at the *beginning* of the paper conveying briefly the content of the paper, and drawing attention to important new information and to the main conclusions. The summary should be intelligible in itself, without reference to the paper, to a reader with some knowledge of the subject; it should not normally exceed 200 words in length. Authors are requested to submit MSS. in duplicate. These should be typed, using double spacing and leaving a margin at one side. Corrections to the MSS. should be made in the text and not in the margin. By Council decision, MSS. of accepted papers are retained by the Society for one year after publication; unless their return is then requested by the author, they are destroyed.

2. *Presentation.*—Authors are allowed considerable latitude, but they are requested to follow the general style and arrangement of *Monthly Notices*. References to literature should be given in the standard form, including a date, for printing either as footnotes or in a numbered list at the end of the paper. Each reference should give the name of the author cited, irrespectively of the occurrence of the name in the text (some latitude being permissible, however, in the case of an author referring to his own work).

3. *Notation.*—Authors should conform closely to the recommendations of Commission 3 of the International Astronomical Union (*Trans. I.A.U.*, Vol. VI, p. 345, 1938). Council has decided to adopt the I.A.U. 4-letter abbreviations for constellations where contraction is desirable (Vol. IV, p. 221, 1932).

4. *Diagrams.*—These should be drawn about twice the size required in print and prepared for direct photographic reproduction except for the lettering which should be inserted in pencil. Legends should be given in the manuscript indicating where in the text the figure should appear. Blocks are retained by the Society for 10 years; unless the author requires them before the end of this period they are then destroyed.

5. *Tables.*—These should be arranged so that they can be printed upright on the page.

6. *Proofs.*—Costs of alteration exceeding 5 per cent of composition must be borne by the author. Fellows are warned that such costs have risen sharply in recent years, and it is in their own and the Society's interests to seek the maximum conciseness and simplification of symbols and equations consistent with clarity.

7. *Revised Manuscripts.*—When papers are submitted in revised form it is especially requested that they be accompanied by the original MS.

Reading of Papers at Meetings

8. If a paper is to be read at a meeting it should be received by the Assistant Secretary not later than the previous Friday.

9. When submitting papers authors are requested to indicate whether they will be willing and able to read the paper at the next or some subsequent meeting, and approximately how long they would like to be allotted for speaking.

10. Postcards giving the programme of each meeting are issued some days before the meeting concerned. Fellows wishing to receive such cards whether for Ordinary Meetings or for the Geophysical Discussions or both should notify the Assistant Secretary.

MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 110 No. 4

MEETING OF 1950 MAY 12

Professor W. M. Smart, President, in the Chair

The election by the Council of the following Fellows was duly confirmed:—

Eric Burgess, 4 Cordova Avenue, Denton, Manchester (proposed by H. L. Dilks);

Bernard Edward Featherstone, M.P.S., 6 Westbourne Road, Sketty, Swansea, Glamorganshire (proposed by G. Corbett);

Rev. Francis J. Heyden, S.J., Ph.D., Georgetown College Observatory, Washington, D.C., U.S.A. (proposed by Rev. P. A. McNally, S.J.); and

Cyril Ashley Swindin, 11 Tyndalls Park Road, Bristol (proposed by F. Sargent).

One hundred and six presents were announced as having been received since the last meeting, including:—

Samuel Cousins, reproduction of engraving of *John Couch Adams* (presented by T. J. Cheater).

The President announced that Professor Joel Stebbins would deliver the George Darwin Lecture at the Ordinary Meeting on 1950 October 13, taking the Electrical Photometry of Stars and Nebulae as his subject.

ADDITIONAL MEETING OF 1950 SEPTEMBER 22

in the Rooms of the Royal Irish Academy, Dublin

Professor W. M. Smart, President, in the Chair

The President referred in the following terms to the sudden death of Professor E. A. Milne (the Fellows standing):

" Fellows of the Society and our guests must now have learned of the severe blow which the Royal Astronomical Society, and indeed the whole of the scientific world, has suffered through the sudden and untimely death of Professor E. A. Milne in Dawson Street yesterday morning while on his way to the rooms of the Academy for our first discussion.

" Many of us have lost a staunch friend and colleague, and the Society has lost a Fellow who must be reckoned as one of the outstanding astronomers of our generation. To an intellect of astonishing range Milne added a clarity and

fluency in exposition and extempore debate which always fascinated his audience. Simple and modest in demeanour, he was always approachable and helpful to Fellows and research students however young in years and experience they might be.

" As a comparatively young man he gained the highest distinction which the Society has to offer—the award of its Gold Medal; and in the anxious years of the war he presided over the fortunes of the Society, steering it successfully through a difficult time.

" His researches cover an extensive field: in the early days at Cambridge, just after the first war, when he and I were colleagues, he solved the problem of the calcium chromosphere; later his researches of far-reaching importance embraced the difficult problems of stellar temperature and stellar structure, and very recently he became deeply interested in the dynamical evolution of the spiral nebulae. All these investigations—and many others—were and are of fundamental importance in astronomy; but perhaps what he himself regarded, as many do, as his most significant contribution to the progress of science, was his original and far-reaching exposition of kinematical relativity—a totally new way of approach to the greatest problem of all, namely the physical properties and evolution of the Cosmos.

" To-day he was to have read an important paper on gravitation and magnetism.

" We have lost a great master whom we in the Royal Astronomical Society will always hold in the highest esteem amongst the giants of our science, and, I must add, in kindly and sincere remembrance."

The President remarked that this meeting was the third in the history of the Society to be held outside London. He expressed the thanks of the Society to the President and Council of the Royal Irish Academy, by whose hospitality the meeting was possible. He also thanked the Minister for External Affairs and Mrs MacBride for their cordial reception of the Society at Iveagh House on the previous evening. The guests, whom the President welcomed, included Mr Eamon de Valera; Mr Geofroy Tory, representing the British Ambassador; Professor M. Tierney, President of University College; the Provost of Trinity College; the Registrar and representatives of the Royal Dublin Society, and of the Dublin Institute of Advanced Studies; the Chairman and members of the Irish Astronomical Society; and the Earl and Countess of Rosse, who had invited Fellows to visit Birr Castle on the following Monday. The President said that the name of the Earl of Rosse was one well known to astronomers and also in the annals of the Society, since two of the present Earl's predecessors had been Fellows. The President remarked that the occasion of the Society's visit to Dublin was of particular interest, since the decision of the Council to hold meetings in centres other than London, for the benefit of Fellows unable to travel so far to meetings, had been prompted by a letter received from a Dublin Fellow, Dr F. E. Lewis, whom he was glad to see present.

Professor F. J. M. Stratton expressed the indebtedness of the Society to Professor H. A. Brück for all that he, and members of his staff, had done to organize the very successful meetings, including the Colloquium on Solar Physics (held the previous day in the rooms of the Royal Irish Academy) and the visit to Dunsink Observatory. His activity showed that Dunsink was once again becoming a centre of astronomical research. A cordial vote of thanks to Professor Brück was passed with acclamation.

Two hundred and forty-four presents were announced as having been received since the last meeting, including :—

W. H. McCrea, *Physics of the Sun and Stars* (presented by the author);

Ernst Zinner, *Leben und Wirken des Johannes Müller von Königsberg, genannt Regiomontanus* (presented by the author);

Paul Couderc, *L'Expansion de l'Univers* (presented by the author);

A. Unsöld, *Physique de l'Atmosphère des Etoiles*—French translation by V. Kourganoff (presented by the Centre National de la Recherche Scientifique); and

Time and its Mysteries, Four lectures given on the James Arthur Foundation, New York University, by D. W. Herring *et alia* (presented by the New York University Press).

A number of Fellows signed the Admission Book. Professor D. H. Menzel showed his latest film, entitled " Action on the Sun ", to illustrate his paper on the classification of solar prominences ; and other papers were read by Professor H. Zanstra, Dr W. H. Ramsey and Mr H. E. Butler.

Other meetings held in connection with the Society's visit to Dublin included, on the evening of September 20, an informal gathering of Fellows in the rooms of the Royal Irish Academy; on the evening of September 22, a Public Lecture on " Time " given by the Astronomer Royal at University College, Dublin, in cooperation with the Dublin Centre of the Irish Astronomical Society; on the morning of September 23, a Geophysical Discussion on Planetary Atmospheres; and on the evening of the same day a reception by the President of University College. Excursions to neighbouring places of interest were arranged for Fellows and their friends; in particular a visit to the Observatory at Armagh on September 24.

GRAVITATION AND MAGNETISM

E. A. Milne

(Received 1950 April 21)

Summary

It is shown by the methods of kinematic relativity that there should be a connection between gravitation and magnetism of the type suggested by the empirical formulae of Blackett and Wilson, multiplied however by certain dimensionless ratios. The field of a rotating system cannot be represented by a dipole. The derivation only applies rigorously to a rotating system like a galaxy, with its centre at a fundamental particle. The results obtained suggest that for a given mean density and given angular momentum, a highly flattened system should have an effective magnetic moment greater than a less flattened system.

---

1. It has been shown by H. W. Babcock (**1**) that the magnetic moments of the Earth, the Sun and the star 78 Virginis are approximately proportional to their respective angular momenta, although since then the work of Thiessen (**2**) has thrown doubt on the reality of the magnetic field of the Sun. The consequences of this proportionality have been developed by Blackett (**3**), who has embodied it in the empirical formula

$$\boldsymbol{\mu} = -\beta \frac{\gamma^{1/2}}{2c} \boldsymbol{U}, \tag{1}$$

where $\boldsymbol{\mu}$ is the magnetic moment of the rotating body, γ the constant of gravitation, c the velocity of light, $\boldsymbol{U}$ the angular momentum and β a constant of the order of unity. Blackett suggested that the empirical relation (1) may be a general property of massive rotating bodies of roughly spherical shape, and that it might indicate a new fundamental property of matter not contained within the structure of existing physical theory. Blackett also pointed out that (1) is equivalent to a hypothesis of H. A. Wilson (**4**), that a mass element m moving with velocity $\boldsymbol{v}$ produces a magnetic field $\boldsymbol{H}$ at a vector distance $\boldsymbol{r}$ given by

$$\boldsymbol{H} = -\beta_1 \frac{\gamma^{1/2}}{c} m \frac{\boldsymbol{v} \wedge \boldsymbol{r}}{|\boldsymbol{r}|^3}. \tag{2}$$

At the same time Blackett emphasized that (2) is certainly untrue if applied to free translation, as it predicts magnetic fields which are not found; and that (2) as it stands is inconsistent with the restricted principle of relativity.

2. It has been claimed by M. Causse (**5**), of the University of Paris, that his development of the theory of kinematic relativity predicts that all forces are of electromagnetic nature, and that this explains the existence of magnetic fields in rotating bodies. I am unable to accept this claim in Causse's form, as it depends on his assumption that inertial mass M is always a constant m, whilst energy E is $mc^2 \log \xi^{1/2}$, a pair of results which do *not* give Einstein's relation $E = Mc^2$, so fundamental in all theories of the ultimate constitution of matter and so devastatingly confirmed by the atomic bomb.

3. In the present paper I show however that a magnetic effect of gravitation can be deduced from the theory of electromagnetism expounded in Part IV of my recent book *Kinematic Relativity* (1948). The point of departure is contained in Section 194 of that work. There I showed that a possible solution of the equations I obtained for the *electromagnetic* super-potentials led to the *gravitational* potential energy for two particles in one another's presence, which I had arrived at by an independent method in Part III of the same work. I am now able to show that the counterpart of the *electric field* corresponding to this solution is precisely the ordinary gravitational field, provided that certain virtual charges are supposed distributed proportional to mass. The question that then arises is: what is the counterpart of the corresponding magnetic field, i. e. the magnetic field arising from the same super-potential? The answer to this question is that the magnetic field is given by H. A. Wilson's formula with the velocity replaced by the cosmical recession velocity corresponding to the place at which the field is being evaluated. After some transformations it is then shown that the effective dipole strength of a body in rotation of the size of the galaxy is given by a formula closely resembling Blackett's formula. Since kinematic relativity was expressly developed to deal with problems of the galactic order of magnitude, it is satisfactory to have got so far. Applied to a body of the size of the Sun, my modification of Blackett's formula gives a negligible magnetic field, but I am not inclined to consider this a legitimate application of the theory, as *prima facie* it only applies to a body in rotation about a fundamental particle in the sense of kinematic relativity, i. e. about the nucleus of a galaxy. Further discussion may be postponed until the formulae have been derived.

4. The electromagnetic theory in kinematic relativity arises from the construction of super-potentials. The super-potential giving the effect at a point P_1 of a charge at a point P_2, all reckoned relative to a given fundamental observer **O**, is a 4-scalar ϕ_{21}, a function of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ and possibly their velocities, such that the 6-vector $(\boldsymbol{H}_1, \boldsymbol{E}_1)$ describing the electromagnetic field at P_1 is given by the relations

$$\boldsymbol{E}_1 = \frac{1}{c}\left[\frac{\partial^2}{\partial \boldsymbol{P}_1 \partial t_2} - \frac{\partial^2}{\partial \boldsymbol{P}_2 \partial t_1}\right]\phi_{21}, \tag{3}$$

$$\boldsymbol{H}_1 = \left[\frac{\partial}{\partial \boldsymbol{P}_1} \wedge \frac{\partial}{\partial \boldsymbol{P}_2}\right]\phi_{21}. \tag{3'}*$$

It has been shown that the super-potentials ϕ_{21} and ϕ_{12} satisfy the equations

$$-\frac{q_1}{ct_0} L_{12}\phi_{21} = -\frac{q_2}{ct_0} L_{21}\phi_{12} = \Phi \tag{4}$$

and

$$L_{11}\phi_{21} = -\phi_{21}, \quad L_{22}\phi_{12} = -\phi_{12}, \tag{5}$$

where

$$L_{\alpha\beta} = t_\alpha \frac{\partial}{\partial t_\beta} + \boldsymbol{P}_1 \cdot \frac{\partial}{\partial \boldsymbol{P}_2}, \qquad (\alpha, \beta = 1, 2)$$

and q_1, q_2 are the charges at P_1, P_2 ; Φ is the joint electromagnetic energy of the two charges q_1 and q_2 ; t_1 and t_2 are the epochs at P_1 and P_2, and t_0 is the present age of the universe, all reckoned from the natural origin of time. Relations (4) and (5) are the conditions that the equations of motion of the two charges (which I need not here quote) shall possess an energy integral.

* The notation $\frac{\partial}{\partial \boldsymbol{P}_1} \wedge \frac{\partial}{\partial \boldsymbol{P}_2}$ expresses a differential operator whose x-component is $\frac{\partial}{\partial y_1 \partial z_2} - \frac{\partial}{\partial z_1 \partial y_2}$.

5. Relations (5) show that ϕ_{21}, ϕ_{12} must be homogeneous and of degree -1 in the variables $(\boldsymbol{P}_1, ct_1)$, $(\boldsymbol{P}_2, ct_2)$ respectively, by Euler's theorem on homogeneous functions. Now previous work had thrown up the denominator

$$(X_{12}^2 - X_1 X_2)^{-1/2},$$

where

$$X_1 = t_1^2 - \boldsymbol{P}_1^2/c^2, \quad X_2 = t_2^2 - \boldsymbol{P}_2^2/c^2, \quad X_{12} = t_1 t_2 - \boldsymbol{P}_1 . \boldsymbol{P}_2/c^2, \tag{6}$$

as playing the part of the denominator r^{-1} in ordinary potential theory ; this is of the requisite degree -1 in each of the pairs of variables $(\boldsymbol{P}_1, ct_1)$, $(\boldsymbol{P}_2, ct_2)$; and moreover

$$L_{12}(X_{12}^2 - X_1 X_2)^{-1/2} = 0 = L_{21}(X_{12}^2 - X_1 X_2)^{-1/2},$$

so that $(X_{12}^2 - X_1 X_2)^{-1/2}$ behaves as a constant under either of the linear operators L_{12}, L_{21}. Hence we can put

$$\phi_{21} = q_2 \frac{\psi_{21}}{(X_{12}^2 - X_1 X_2)^{1/2}}, \quad \phi_{12} = q_1 \frac{\psi_{12}}{(X_{12}^2 - X_1 X_2)^{1/2}}, \tag{7}$$

where ψ_{21} must be of degree zero in the variables $(\boldsymbol{P}_1, ct_1)$, ψ_{12} of degree zero in the variables $(\boldsymbol{P}_2, ct_2)$.

It was shown that the solution which gave classical electromagnetism in a strictly relativistic form was

$$\psi_{21} = -\tfrac{1}{2} \frac{Z_2 Z_{21}}{Y_1^{1/2} Y_2^{1/2}}, \quad \psi_{12} = -\tfrac{1}{2} \frac{Z_1 Z_{12}}{Y_1^{1/2} Y_2^{1/2}}, \tag{8), (8'}$$

where

$$Z_1 = t_1 - \boldsymbol{P}_1 . \boldsymbol{V}_1/c^2, \quad Z_2 = t_2 - \boldsymbol{P}_2 . \boldsymbol{V}_2/c^2,$$

$$Z_{12} = t_1 - \boldsymbol{P}_1 . \boldsymbol{V}_2/c^2, \quad Z_{21} = t_2 - \boldsymbol{P}_2 . \boldsymbol{V}_1/c^2, \tag{9), (9'}$$

$$Y_1 = 1 - \boldsymbol{V}_1^2/c^2, \qquad Y_2 = 1 - \boldsymbol{V}_2^2/c^2, \tag{10), (10'}$$

$\boldsymbol{V}_1$, $\boldsymbol{V}_2$ being the velocities of the charges ρ_1 and ρ_2 at P_1 and P_2. (I am using the symbol ρ to denote actual charges, the symbol q to denote virtual charges.)

6. On the other hand, I pointed out in Section 194 of *Kinematic Relativity* that a solution *not* involving the velocities was

$$\psi_{21} = X_2, \quad \psi_{12} = X_1. \tag{11), (11'}$$

For then

$$L_{12}\psi_{21} = 2X_{12} = L_{21}\psi_{12}. \tag{12}$$

We have then

$$\phi_{21} = q_2 \frac{X_2}{(X_{12}^2 - X_1 X_2)^{1/2}}, \quad \phi_{12} = q_1 \frac{X_1}{(X_{12}^2 - X_1 X_2)^{1/2}}, \tag{13}$$

whence, by (4),

$$\Phi = -\frac{q_1 q_2}{ct_0} \frac{2X_{12}}{(X_{12}^2 - X_1 X_2)^{1/2}}. \tag{14}$$

7. Now the joint *gravitational* energy of two particles of masses m_1, m_2 at points P_1, P_2 at times t_1, t_2 had been shown by independent reasoning to be χ, given by

$$\chi = -\frac{m_1 m_2 c^2}{M_0} \frac{X_{12}}{(X_{12}^2 - X_1 X_2)^{1/2}}, \tag{15}$$

where M_0 is the mass of the equivalent homogeneous universe. Let us now, as a definite physical hypothesis suggested by the formal equivalence of (14) and (15),

identify the pseudo-electromagnetic energy Φ with the gravitational energy χ, by assigning to each particle of mass m a virtual charge q, proportional to mass, given by

$$\frac{2q^2}{ct_0} = \frac{m^2c^2}{M_0}, \tag{16}$$

or

$$q\sqrt{2} = \gamma_0^{1/2} m, \tag{17}$$

where γ_0, given by

$$\gamma_0 = \frac{c^3 t_0}{M_0}, \tag{18}$$

is the present value of the so-called constant of gravitation. That this is the correct interpretation of γ_0 is shown by taking the observer at P_2 and so putting $\boldsymbol{P}_2 = 0$ in (15), when we get

$$\chi = -\frac{m_1 m_2 c^2}{M_0}\frac{ct_1}{|\boldsymbol{P}_1|} = -\frac{\gamma_1 m_1 m_2}{|\boldsymbol{P}_1|}, \tag{19}$$

where

$$\gamma_1 = c^3 t_1 / M_0, \tag{19'}$$

t_1 being the epoch at which P_1 is considered. It is well known that (18) reproduces the actual value of γ_0, on taking for t_0 the present age of the universe as judged by the recession of the extragalactic nebulae, $0{\cdot}6 \times 10^{17}$ sec., and for M_0 the value $2{\cdot}6 \times 10^{55}$ g. as corresponding to a present mean density of the smoothed-out universe of 10^{-27} g. cm.$^{-3}$. The proportionality $q^2 \propto t_0$ is consistent with the properties of the fine-structure constant $2\pi e^2/hc$, since this is dimensionless and since kinematic relativity requires $h \propto t$.

The present stage of our calculation is then that we have identified the joint gravitational energy of two massive particles m_1, m_2 with the joint electromagnetic energy of two virtual point-charges q_1, q_2. This virtual electromagnetic field is supposed to be additional to any actual electromagnetic field due to any actual charges e that may be present. Let us now calculate the corresponding virtual electric intensity and compare it with the gravitational attraction.

8. The gravitational attraction at the test-particle P_1, of mass m_1, due to the mass m_1 at P_2, was shown to be given in Lorentz-invariant form by

$$-\frac{\partial \chi}{\partial \boldsymbol{P}_1} = \frac{m_1 m_2}{M_0}\frac{X_2}{(X_{12}^2 - X_1 X_2)^{3/2}}(\boldsymbol{P}_2 X_1 - \boldsymbol{P}_1 X_{12}). \tag{20}$$

The virtual electric intensity at P_1 due to q_2 at P_2 is, on the other hand, given by

$$\boldsymbol{E}_1 = \frac{1}{c}\left[\frac{\partial}{\partial \boldsymbol{P}_1}\frac{\partial}{\partial t_2} - \frac{\partial}{\partial \boldsymbol{P}_2}\frac{\partial}{\partial t_1}\right]\frac{q_2 X_2}{(X_{12}^2 - X_1 X_2)^{1/2}}. \tag{21}$$

Using the relations

$$\frac{\partial^2}{\partial \boldsymbol{P}_1 \partial t_2} X_2 \equiv 0 \equiv \frac{\partial^2}{\partial \boldsymbol{P}_2 \partial t_1} X_2, \tag{22}$$

$$\left[\frac{\partial^2}{\partial \boldsymbol{P}_1 \partial t_2} - \frac{\partial^2}{\partial \boldsymbol{P}_2 \partial t_1}\right]\frac{1}{(X_{12}^2 - X_1 X_2)^{1/2}} \equiv 0, \tag{23}$$

we find that threatened terms with denominators $(X_{12}^2 - X_1 X_2)^{-1/2}$, $(X_{12}^2 - X_1 X_2)^{-5/2}$ disappear identically, and we get by detailed differentiation and some cancellation of terms,

$$\boldsymbol{E}_1 = \frac{2q_2}{c^3}\frac{X_2}{(X_{12}^2 - X_1 X_2)^{3/2}}(t_1 \boldsymbol{P}_2 - t_2 \boldsymbol{P}_1). \tag{24}$$

For $\boldsymbol{P}_2=0$, (20) reduces to

$$-\frac{\partial\chi}{\partial\boldsymbol{P}_1}=-\left(\frac{m_1m_2c^3t_1}{M_0}\right)\frac{\boldsymbol{P}_1}{|\boldsymbol{P}_1|^3}, \tag{20'}$$

and (24) gives a pseudo-electric force on q_1 equal to $q_1\boldsymbol{E}_1$ or

$$-2q_1q_2\frac{\boldsymbol{P}_1}{|\boldsymbol{P}_1|^3}. \tag{24'}$$

On using (16) we see that (20′) and (24′) are identical, and give moreover an inverse square law of force. Furthermore, (24′) represents an ***attraction*** on P_1 in the direction of P_2. Thus (21) gives the correct sign for a gravitation intensity. That this is no trivial point is seen by taking the genuine electromagnetic solution (8), when we find (Section 195 of *Kinematic Relativity*)

$$\boldsymbol{E}_1=\frac{e_2}{Y_1^{\frac{1}{2}2}Y_2^{\frac{1}{2}2}}\frac{(\boldsymbol{P}_1X_2-\boldsymbol{P}_2X_{12})^{\frac{1}{2}}(Z_2+Z_{21})-(t_1X_2-t_2X_{12})^{\frac{1}{2}}(\boldsymbol{V}_1Z_{12}+\boldsymbol{V}_2Z_{21})}{(X_{12}^2-X_1X_2)^{3/2}}, \tag{25}$$

which gives, on taking the observer at P_2 ($\boldsymbol{P}_2=0$),

$$e_1\boldsymbol{E}_1=+\frac{e_1e_2}{Y_1^{1/2}Y_2^{1/2}}\frac{\boldsymbol{P}_1}{|\boldsymbol{P}_1|^3}. \tag{25'}$$

Thus " like " actual charges repel, " like " virtual charges attract, as required by experience. It should be noted that the form (24), like (25), represents the electric intensity as part of a tensor, namely $\boldsymbol{E}_1$ as a component of a 6-vector, whilst (20) represents the gravitational attraction as a 3-vector.

9. The question now arises, if the virtual ***electric*** intensity corresponds to gravitational attraction, what corresponds to the virtual ***magnetic*** intensity? This is given by

$$\boldsymbol{H}_1=\left[\frac{\partial}{\partial\boldsymbol{P}_1}\wedge\frac{\partial}{\partial\boldsymbol{P}_2}\right]\frac{q_2X_2}{(X_{12}^2-X_1X_2)^{1/2}}. \tag{26}$$

Carrying out the actual differentiation and noting that

$$\left[\frac{\partial}{\partial\boldsymbol{P}_1}\wedge\frac{\partial}{\partial\boldsymbol{P}_2}\right]X_2\equiv 0, \tag{27}$$

$$\left[\frac{\partial}{\partial\boldsymbol{P}_1}\wedge\frac{\partial}{\partial\boldsymbol{P}_2}\right]\frac{1}{(X_{12}^2-X_1X_2)^{1/2}}\equiv 0, \tag{27'}$$

we find

$$\boldsymbol{H}_1=\frac{2q_2}{c^4}\frac{X_2}{(X_{12}^2-X_1X_2)^{3/2}}\boldsymbol{P}_1\wedge\boldsymbol{P}_2. \tag{28}$$

Taking $t_1=t_2=t$, this gives approximately

$$\boldsymbol{H}_1=2q_2\frac{\boldsymbol{P}_1}{ct}\wedge\frac{\boldsymbol{P}_2}{|\boldsymbol{P}_2-\boldsymbol{P}_1|^3}, \tag{29}$$

which may be re-written as

$$\boldsymbol{H}_1=2q_2\frac{\boldsymbol{P}_1}{ct}\wedge\frac{\boldsymbol{P}_2-\boldsymbol{P}_1}{|\boldsymbol{P}_2-\boldsymbol{P}_1|^3}. \tag{30}$$

This is the moment about the origin of a vector through P_1 in the direction towards P_2 inversely proportional to the square of the distance $|\boldsymbol{P}_2-\boldsymbol{P}_1|$. This has no counterpart in classical gravitational theory.

10. If we re-write (30) as

$$\boldsymbol{H}_1=-2q_2\frac{\boldsymbol{P}_2}{ct}\wedge\frac{\boldsymbol{P}_1-\boldsymbol{P}_2}{|\boldsymbol{P}_1-\boldsymbol{P}_2|^3}, \tag{31}$$

and note that $\boldsymbol{P}_2/t$ is the cosmical recession velocity $\boldsymbol{V}_2$ at the source-point P_2, we see that (31) is just Wilson's formula for this case; for putting $\boldsymbol{P}_1-\boldsymbol{P}_2=\boldsymbol{r}$, it gives

$$\boldsymbol{H}_1=-2q_2\frac{\boldsymbol{V}_2}{c}\wedge\frac{\boldsymbol{r}}{|\boldsymbol{r}|^3}, \tag{32}$$

or, using (17),

$$\boldsymbol{H}_1=-\frac{(2\gamma)^{1/2}}{c}m_2\boldsymbol{V}_2\wedge\frac{\boldsymbol{r}}{|\boldsymbol{r}|^3}, \tag{32a}$$

which is (2) with β_1 evaluated as $\sqrt{2}$.

11. It must be particularly emphasized that we have not established H. A. Wilson's formula as true generally. Indeed, had we done so, we should have established too much, for, as emphasized by Blackett and Wilson, the formula cannot be true in general. By obtaining (28) by processes strictly relativistic we have however removed the reproach which Blackett made against Wilson's formula, that it was not consistent with relativity. Thus an appropriate modification of Wilson's formula is established in a strictly relativistic form when $\boldsymbol{V}_2$ is the cosmical recession velocity at P_2 with respect to the observer, taken to be at a fundamental particle, provided we regard the virtual magnetic field as an actual magnetic field.

12. Let us now consider the bearing of this on the rotation of a massive system of particles centred on a fundamental particle. Consider a particle m_1 of such a system, at a point P_1. The gravitational force $\boldsymbol{F}_1$ on m_1 due to all the other particles in the universe is given by

$$\boldsymbol{F}_1=\gamma m_1\sum_{s\neq 1}m_s\frac{\boldsymbol{P}_s-\boldsymbol{P}_1}{|\boldsymbol{P}_s-\boldsymbol{P}_1|^3}, \tag{33}$$

and the moment $\boldsymbol{\Gamma}_1$ of this force about the origin is given by*

$$\boldsymbol{\Gamma}_1=\gamma m_1\boldsymbol{P}_1\wedge\sum_{s\neq 1}m_s\frac{\boldsymbol{P}_s-\boldsymbol{P}_1}{|\boldsymbol{P}_s-\boldsymbol{P}_1|^3}. \tag{34}$$

We can express this in terms of $\boldsymbol{H}_1$, by (30); thus, since now by (30)

$$\boldsymbol{H}_1=\frac{2\boldsymbol{P}_1}{ct}\wedge\sum_{s\neq 1}q_s\frac{\boldsymbol{P}_s-\boldsymbol{P}_1}{|\boldsymbol{P}_s-\boldsymbol{P}_1|^3}, \tag{35}$$

we have

$$\boldsymbol{\Gamma}_1=\gamma m_1\frac{ct}{2}\frac{m}{q}\boldsymbol{H}_1,$$

or, by (17),

$$\boldsymbol{\Gamma}_1=(\tfrac{1}{2}\gamma)^{1/2}ct\,m_1\boldsymbol{H}_1. \tag{36}$$

Now it is a well-established result in kinematic relativity that the angular momentum of a particle or system of particles about a fundamental particle increases secularly proportionally to the time reckoned from the natural origin of time, this effect being due to the couple exerted by the substratum or smoothed-out universe. The angular momentum $\boldsymbol{U}$ obeys in fact the equation (*Kinematic Relativity*, Section 98)

$$\frac{d\boldsymbol{U}}{dt}=\frac{\boldsymbol{U}}{t}. \tag{37}$$

---

* The contribution to $\boldsymbol{\Gamma}_1$ of the *mechanical force* due to $\boldsymbol{H}_1$ is zero; for it is

$$q_1\boldsymbol{P}_1\wedge\left(\frac{\boldsymbol{v}_1\wedge\boldsymbol{H}}{c}\right)=q_1[-\boldsymbol{H}_1(\boldsymbol{P}_1\cdot\boldsymbol{v}_1)+\boldsymbol{v}_1(\boldsymbol{P}_1\cdot\boldsymbol{H}_1)].$$

But $\boldsymbol{P}_1\cdot\boldsymbol{H}_1=0$, by (28); and for a rotating body, for which $\boldsymbol{v}_1=\boldsymbol{\Omega}\wedge\boldsymbol{P}_1$, we have also $\boldsymbol{P}_1\cdot\boldsymbol{v}_1=0$.

But we may equally calculate $d\boldsymbol{U}/dt$ as equal to the vector sum of the couples exerted by the separate gravitational pulls of the remaining particles of the universe. Thus we may put

$$\frac{d\boldsymbol{U}}{dt} = \sum_1 \boldsymbol{\Gamma}_1, \tag{38}$$

where $\sum_1$ denotes a summation over all the particles of the system, of which a particular particle was m_1. Hence, combining (36), (37) and (38), we have

$$\frac{\boldsymbol{U}}{t} = (\tfrac{1}{2}\gamma)^{1/2} ct \sum_1 m_1 \boldsymbol{H}_1. \tag{39}$$

We now give t its present value t_0, and consequently γ its present value γ_0. Then

$$\sum_1 m_1 = \boldsymbol{H}_1 = \frac{1}{ct_0^2} \frac{1}{(\tfrac{1}{2}\gamma_0)^{1/2}} \boldsymbol{U}. \tag{40}$$

13. The virtual magnetic field $\boldsymbol{H}_1$ at P_1, given by (35), is not that due to a dipole at the origin. The magnetic field $\boldsymbol{H}_1$ at P_1 due to a dipole of moment $\boldsymbol{\mu}$ at the origin is in fact given by

$$\boldsymbol{H}_1 = -\frac{\partial}{\partial \boldsymbol{P}_1} \frac{\boldsymbol{\mu}.\boldsymbol{P}_1}{|\boldsymbol{P}_1|^3} = -\frac{\boldsymbol{\mu}}{|\boldsymbol{P}_1|^3} + \frac{3(\boldsymbol{\mu}.\boldsymbol{P}_1)\boldsymbol{P}_1}{|\boldsymbol{P}_1|^5}, \tag{41}$$

and we can assign no value to $\boldsymbol{\mu}$ independent of P_1 which makes (41) reproduce (35). Nevertheless, since Babcock and Blackett have conducted their analyses on the hypothesis that the magnetic field of a large rotating body is that due to a dipole parallel to the axis of rotation, it is convenient to pick on the value of $\boldsymbol{H}_1$ at, say, points on the equatorial plane of the rotating body, and represent the value as due to a dipole of moment $\boldsymbol{\mu}_{eff}$ parallel to the axis of rotation. Then, since for a point P_1 at the equator, $\boldsymbol{\mu}_1 . \boldsymbol{P}_1 = 0$, we have by (41)

$$\boldsymbol{H}_1 = -\frac{\boldsymbol{\mu}_{eff}}{|\boldsymbol{P}_1|^3}. \tag{42}$$

This value $\boldsymbol{H}_1$ takes the same value at all points on the equator, and if a is the equatorial radius,

$$\boldsymbol{H}_1 = -\frac{\boldsymbol{\mu}_{eff}}{a^3}. \tag{43}$$

Hence, with this convention, the summation occurring on the left-hand side of (40), when extended to a set of particles forming an equatorial ring of mass (Σm_1) and radius a, comes to

$$\sum_1 m_1 \boldsymbol{H}_1 = -\frac{\boldsymbol{\mu}_{eff}}{a^3} \sum_1 m_1, \tag{44}$$

and so, by (40),

$$\boldsymbol{\mu}_{eff} = -\frac{1}{ct_0^2} \frac{1}{(\tfrac{1}{2}\gamma_0)^{1/2}} \frac{a^3 \boldsymbol{U}}{\sum_1 m_1}. \tag{45}$$

Multiply numerator and denominator on the right-hand side by $\gamma_0^{1/2}$, and use for γ_0 in the denominator the value given by (18). We get

$$\boldsymbol{\mu}_{eff} = -\sqrt{2} \frac{\gamma_0^{1/2}}{c} \frac{M_0}{\frac{4}{3}\pi(ct_0)^3} \frac{\frac{4}{3}\pi a^3 \boldsymbol{U}}{\sum_1 m_1}. \tag{46}$$

Now $M_0/\frac{4}{3}\pi(ct_0)^3$ is the mean density of the equivalent homogeneous universe, since in kinematic relativity ct_0 is the radius of the universe. Also $\boldsymbol{U}$, being the

angular momentum of the ring of particles of radius a and mass Σm_1, is equal to $(\Sigma m_1)\Omega a^2$, $\boldsymbol{\Omega}$ being the vector angular velocity. Thus [1]

$$\mu_{eff} = -\sqrt{2}\frac{\gamma_0^{1/2}}{c}\rho_0 \tfrac{4}{3}\pi a^5 \boldsymbol{\Omega}. \tag{47}$$

14. Next, suppose that the system of particles we are considering forms part of an ellipsoidal galaxy of equatorial radii a and polar semi-axis b. For such a system, regarded as homogeneous, the *total* angular momentum $\boldsymbol{A}$ is equal to $\frac{2}{5}Ma^2\boldsymbol{\Omega}$, and the *total* mass M is equal to $\frac{4}{3}\pi\rho_1 a^2 b$, where ρ_1 is the mean density of the galaxy. In terms of $\boldsymbol{V}$ and ρ_1, we have

$$\tfrac{4}{3}\pi a^5 \boldsymbol{\Omega} = \frac{a}{b}\frac{\frac{5}{2}\boldsymbol{A}}{\rho_1}. \tag{48}$$

Hence (47) gives

$$\mu_{eff} = -\tfrac{5}{2}\sqrt{2}\frac{\gamma_0^{1/2}}{c}\frac{\rho_0}{\rho_1}\frac{a}{b}\boldsymbol{A}. \tag{49}$$

Formula (49) gives the effective magnetic moment of a rotating galaxy in the form of Blackett's formula multiplied by the dimensionless factors ρ_0/ρ_1 and a/b.

It is evident that these factors act in opposite directions; the factor ρ_0/ρ_1 is considerably less than unity, the factor a/b considerably greater than unity. The mean density of the smoothed-out universe, ρ_0, cannot be greater than 10^{-27} g. cm.$^{-3}$; the mean density ρ_1 of our own galaxy cannot be less than 10^{-26} g. cm.$^{-3}$. Hence the factor ρ_0/ρ_1 cannot exceed $\frac{1}{10}$ at most, and is more likely to be at least as small as $\frac{1}{100}$. The factor a/b, on the other hand, may be taken to be at least as great as 10, possibly much larger for a discoidal galaxy. We may be satisfied with having established theoretically the same order of magnitude for the magnetic moment of a rotating galaxy as was suggested by Blackett's empirical formula.

15. Formula (49) indicates that extreme flattening, for a given mean density, is highly favourable to the existence of a large magnetic dipole moment, in the case of a galaxy. Great caution must be used, on the other hand, in attempting to apply (49) to smaller systems, whose centre is not at a fundamental particle. *Prima facie*, for a body like the Sun, for which ρ_1 is of the order of magnitude of unity and a/b equal to unity, formula (49) predicts a negligible magnetic moment in so far as such a moment could arise from the virtual charge-distribution considered in the present paper. But the Sun is not centred at a fundamental particle, in the sense of kinematic relativity.

16. It may be repeated that the form of field predicted by the present theory is not that of a dipole. If we apply to (40) the transformations used in the immediately succeeding analysis, but *without* deriving $\boldsymbol{H}_1$ from a dipole, we get for the equatorial value of the magnetic field at distance a the more fundamental formula

$$\boldsymbol{H}_1 = \sqrt{2}\frac{\gamma_0^{1/2}}{c}\rho_0 \tfrac{4}{3}\pi a^2 \boldsymbol{\Omega}, \tag{50}$$

a result which does not assume an ellipsoidal form for the galaxy or its homogeneity.

17. Let us apply (50) to our own galaxy. Taking

$$\gamma_0 = 6{\cdot}66 \times 10^{-8}, \quad c = 3 \times 10^{10}\,\text{cm. sec.}^{-1}, \quad \rho_0 = 10^{-27}\,\text{g. cm.}^{-3},$$

$\boldsymbol{\Omega}$ corresponding to a rotation period at the Sun's distance from the centre of $2{\cdot}24 \times 10^8$ years, and $a = 10^4$ parsec $= 3{\cdot}08 \times 10^{22}$ cm., we find from (50) that the galactic magnetic field at the Sun's distance should be $0{\cdot}45 \times 10^{-10}$ gauss.

Professor H. H. Plaskett has kindly drawn my attention to a paper by W. A. Hiltner (**6**), wherein observed polarization effects in the light from distant stars

are attributed to possible scattering by interstellar particles oriented by a magnetic field. Such a magnetic field has been postulated on other grounds by Alfvén (**7**) and evaluated as of the order of 10^{-10} gauss, which is about what is required to make the shower particles in cosmic rays fairly isotropic. It should be noted that (50) contains no arbitrary coefficients, and does not require an estimate of the mass of the galaxy.

18. The most fundamental result of the present paper is (35) with q_s replaced by m_s by use of (17), namely,

$$\boldsymbol{H}_1 = \sqrt{2}\frac{\gamma_0^{1/2}}{ct_0}\boldsymbol{P}_1 \wedge \sum_{s\neq 1} m_s \frac{\boldsymbol{P}_s - \boldsymbol{P}_1}{|\boldsymbol{P}_s - \boldsymbol{P}_1|^3}. \tag{51}$$

We may state this new fundamental property of matter in words in the form that the gravitational magnetic field at any point P_1 is proportional to the resultant moment about the origin (taken as a fundamental particle) of all the gravitational forces at P_1 due to the rest of the universe.

Oxford:
1950 *April* 20.

Notes added 1950 *June* 17

1. (Note to Section 4.) The epochs t_1, t_2 at P_1, P_2 are to be treated as independent variables in differentiating the super-potentials. When application is made to an actual electromagnetic or gravitational situation, an invariant relation connecting t_1 and t_2 needs to be imposed. This may be taken in the gravitational case to be $X_1 = X_2$. (See Camm (**8**).) This reduces for all ordinary distances to $t_2 = t_1$, approximately. The same relation may be imposed in electromagnetic cases when radiation effects are not involved. However, in many applications of the present electromagnetic theory the explicit occurrence of two epochs t_1, t_2 is merely formal, as on taking the observer at P_2, so that $\boldsymbol{P}_2 = 0$, the epoch t_2 is found to disappear. (*Cf.* (15) and (19), and also (25) and (25′).) We need then only the value of t_1 at the test particle, and this is just the current time-variable as used by the observer at the origin.

2. (Note to Section 11.) In the text, the virtual magnetic field is regarded as an actual magnetic field, acting on any moving charge, whereas the electric field, identified with gravitation, acts only on the virtual electric charge q_1. I was aware of this inconsistency, and a referee has invited me to justify it. I think that if the magnetic field is to be regarded as a new fundamental property of matter, we shall just have to accept this as a part of the strangeness of the situation. On the other hand, it is possible that we should invert the relation, and regard the gravitational pulls on gross matter as actual electric forces acting on real charges. It may be pointed out that the virtual charge q which I have associated with a proton of mass $m = 1{\cdot}66 \times 10^{-24}$ g., according to (17), is only $3{\cdot}0 \times 10^{-28}$ e.s.u., as compared with the actual charge on a proton of $4{\cdot}8 \times 10^{-10}$ e.s.u.

References

(1) Babcock, H. W., *Astrophys. Journ.*, **105**, 105, 1947.
(2) See Babcock, H. W., *P.A.S.P.*, **60**, 244, 1948.
(3) Blackett, P. M. S., *Nature*, **169**, 658, 1947; *Phil. Mag.*, ser. 7, **40**, 125, 1949.
(4) Wilson, H. A., *Proc. Roy. Soc.* A, **104**, 451, 1923.
(5) Causse, M., *Thèse, Fac. Sci. de l'Univ. de Paris*, 26ff., 1949; *Ann. de Phys.*, ser. 12, **4**, 760, 1949.
(6) Hiltner, W. A., *Astrophys. Journ.*, **109**, 478, 1949.
(7) Alfvén, H., *Cosmical Electrodynamics* (Clarendon Press), 1950, p. 223.
(8) Camm, G. L., *Nature*, **155**, 754, 1945.

MODELS FOR RED GIANT STARS

I. General Discussion and Application to Homogeneous Models

C. M. Bondi

(Received in revised form 1950 May 1*)

Summary

The aim of this series of papers is to consider systematically all possible models for red giant stars from the mathematical point of view. This first paper gives a brief general discussion of the red giant problem and this is followed by a consideration of the mathematically simplest case, that of homogeneous models. It is shown that if present views on opacity laws in stars are correct, no homogeneous model provides an adequate explanation of the red giant phenomenon.

---

1. *Introduction.*—The problem of red giant stars has received much attention from many writers, but has so far eluded any complete solution. Various models have been suggested, some of which give considerable increases in radius over a main sequence star of the same mass, but there has been no systematic investigation to discover all the possibilities and then to select those models which have some physical relevance.

Red giant stars as observed have luminosities equal approximately to some of the highest luminosities in the main sequence, but have very much lower surface temperatures. This fact indicates that the radii must be vastly greater than those of main sequence stars with these fairly high luminosities. As far as the masses are concerned, they are not very well known, but what is known shows that they do not differ very greatly from the higher main sequence masses. The type of phenomenon for which a theory must account is the existence of stars of approximately the same mass but with a radius of about 200 times that of a main sequence star.

The law of energy generation applicable in red giant stars requires some comment. The possibility that red giants might be stars of low central temperature producing their energy by gravitational contraction can be ruled out as a complete explanation on the grounds that the lifetime would be much too short to account for the proportion of red giants observed. Hence the energy must be supplied by nuclear processes, which requires that the central temperature should be in the neighbourhood of 2×10^7 degrees. The central temperature must therefore be approximately the same as that of a main sequence star.

It is the aim of this paper to begin a systematic investigation of the problem and in particular to deal with homogeneous models. It is shown that by the use of h-variables described in an earlier paper†, the criterion of a large increase in radius over a main sequence star of the same mass is very simply expressible. When a solution of a stellar model is plotted in the (S, Q) plane, a qualitative

* Originally submitted 1950 January 10.

† C. M. Bondi and H. Bondi, *M.N.* **109**, 62, 1949. (To be referred to as Paper A.)

examination shows at once whether or not a large increase in radius can be expected for that model. Thus it is necessary to integrate only those solutions which appear to give results of interest.

The models considered in this paper belong to the general class of those with a convective core and an atmosphere in radiative equilibrium. This type of model is widely accepted as being the only valid one when nuclear processes are responsible for the energy production, but in fact this choice has very little effect on the radius. The polytropic index for the core is taken to be 1·5. Radiation pressure is neglected throughout. This cannot have a major effect on the radius and seems justifiable at this stage.

2. In this paper, the variables r, ρ, P and T have their usual definitions. Other symbols are defined as follows:—

$M = M(r)$ = mass contained within a radius r of the centre of the star,

G = constant of gravitation,

$\mathscr{R}$ = gas constant,

μ = molecular weight,

a = Stefan's constant,

c = speed of light,

X = proportion of hydrogen by mass,

$S = 4\pi Pr^4/GM^2$,

$Q = Pr/GM\rho$,

$N = 1 - d\log\rho/d\log P$,

n = polytropic index as it occurs in Emden's equation,

$Q_E(S; N)$ = Emden function corresponding to $N = 1/n + 1$.

Suffix s will be used for surface values and suffix c for central values.

S, Q, N are the h-variables as defined in Paper **A**.

The hydrostatic equation in physical variables is

$$\frac{dP}{dr} = -\frac{GM\rho}{r^2}, \tag{1}$$

and the equation of mass conservation is

$$\frac{dM}{dr} = 4\pi r^2\rho. \tag{2}$$

When these are expressed in h-variables, they lead to the single equation

$$\frac{S}{Q}\frac{dQ}{dS} = \frac{S+N-Q}{1+2S-4Q}. \tag{3}$$

The variables r/r_s, M/M_s and T/T_c can be considered as functions of S and obtained by the following quadratures.

$$\log\frac{r_s}{r} = \int_0^S \frac{Q\,dS}{S(1+2S-4Q)} = \int_0^Q \frac{dQ}{S+N-Q},$$

$$\log\frac{M_s}{M} = \int_0^Q \frac{S\,dQ}{Q(S+N-Q)} = \int_0^S \frac{dS}{1+2S-4Q},$$

$$\log\frac{T_c}{T} = \int_S^\infty \frac{N\,dS}{S(1+2S-4Q)} = \int_Q^\infty \frac{N\,dQ}{Q(S+N-Q)}.$$

In addition to equation (3) we have the energy transfer equation which takes various forms under various physical conditions. In almost all cases, this can be expressed in terms of h-variables.

Thus one can obtain solutions in the form $Q=Q(S)$ and $N=N(S)$.

In comparing the radii of different stars, it is of great advantage to have available some non-dimensional quantity which is a measure of the radius of a star, given its mass. For, even two main sequence stars may have very different radii if the ratio of their masses is large. It is the fact that their radii are so much larger than the radii of main sequence stars of the same mass that is the phenomenon of red giants. In constructing a non-dimensional parameter, the central temperature is available, since it cannot vary much among stars which generate their energy by nuclear processes, and hence must be nearly the same for red giants as for main sequence stars.

The invariant

$$\mathscr{R}T_c r_s/\mu_c G M_s$$

possesses the required qualities. The factor μ_c in the denominator can contribute only a small factor. We therefore consider the expression of this invariant in terms of h-variables.

From the definition of Q, it follows that

$$\log Q=\log\frac{Pr}{GM\rho}=\log\frac{P_c r_s}{GM_s\rho_c}+\log\frac{r}{r_s}+\log\frac{P}{P_c}-\log\frac{M}{M_s}-\log\frac{\rho}{\rho_c}.$$

Therefore, since $P_c=\mathscr{R}\rho_c T_c/\mu_c$,

$$\log\frac{\mathscr{R}T_c r_s}{\mu_c G M_s}=\log Q_1+\int_0^{Q_1}\frac{dQ}{S+N-Q}-\int_0^{S_1}\frac{dS}{1+2S-4Q}+\int_{S_1}^{\infty}\frac{N\,dS}{S(1+2S-4Q)}$$

from the quadratures for r/r_s, M/M_s and T/T_c, where (S_1, Q_1) is an *arbitrary* point on the (S, Q) solution. It must be pointed out that the integrals in the above expression are line integrals, taken along the particular solution under consideration.

That the right-hand side of the above equation is independent of the particular point (S_1, Q_1) may be seen by differentiation with respect to S_1, using equation (3).

$\mathscr{R}T_c r_s/\mu_c G M_s$ is an invariant for all stars possessing the same homology invariant solution. In general, this is a single infinity of stars.

Suppose that a specific value of S_1 is taken, sufficiently large for it to be a point of the convective core for all solutions under consideration. That this is legitimate follows from the arbitrariness of S_1, discussed above. We now compare the values of the invariant for a standard main sequence model and for an extended model, using the same value of S_1 in each case. Both $\log Q_1$ and

$$\int_{S_1}^{\infty}\frac{N\,dS}{S(1+2S-4Q)}$$

are independent of the particular solution, since S_1 lies in the core for each model. Hence large extensions of radius will occur when

$$\log\frac{r_s}{r_1}=\int_0^{Q_1}\frac{dQ}{(S+N-Q)}=\int_0^{S_1}\frac{dQ}{dS}\frac{dS}{(S+N-Q)}$$

(where suffix 1 indicates values at S_1) is much larger than the corresponding quantity for a main sequence star, provided that at the same time

$$\log\frac{M_s}{M_1}=\int_0^{S_1}\frac{dS}{(1+2S-4Q)}$$

does not become large.

r_1, M_1 are dimensional quantities and are not necessarily the same for two different models at the same point (S_1, Q_1). Hence r_1, M_1 are not useful standards of comparison for r_s, M_s and $\log r_s/r_1$, $\log M_s/M_1$ are not in themselves appropriate criteria for red giant stars. It is only as a consequence of their appearance in the expression for the invariant that they are of interest.

From the second integral form for r_s/r_1, it follows that in regions where $S+N-Q$ is small, but where dQ/dS is not very small, the contribution of this integral to the invariant is large. From the integral form for M_s/M_1, it follows that in order that the contribution of $\log M_s/M_1$ to the invariant shall not be too large, $(1+2S-4Q)$ must not be very small. Thus the invariant will be large if over some part of the solution dQ/dS and $(1+2S-4Q)$ are of the order of magnitude of unity, while $S+N-Q$ is small compared with unity.

Now, from equation (3) it follows that in order that the condition on dQ/dS shall be compatible with that on $(S+N-Q)$, either S or $(1+2S-4Q)$ must be small. As a consequence of the above condition on $(1+2S-4Q)$, it follows that S must be small compared with unity.

The only possible means by which r_s/r_1 may be made large without making $S+N-Q$ small is by making dQ/dS large. This would imply again that S should be small compared with unity, since $(1+2S-4Q)$ is of the order of magnitude of unity as before.

The only points at which dQ/dS is large, while $(S+N-Q)$ and $(1+2S-4Q)$ are of order unity, are in the neighbourhood of $S=Q=0$, that is, near the surface. This is a region through which the solutions for all models must pass. Hence the region near to $S=Q=0$ makes a large contribution to the invariant whether the standard model is used or not. Since the aim is to produce an increase of the invariant over its value for a main sequence star, this region is therefore not relevant.

It is therefore both necessary and sufficient, in order that a stellar model may possess a value of the invariant large compared with its value for the standard model, that the solution should pass through a region in which both S and $S+N-Q$ are small compared with unity. It follows that the solution must pass through a region in which both S and $Q-N$ are small compared with unity.

But for small S, N is always a slowly varying function of S. Hence, when S is small, $N-N_s$ is also small and, therefore, the above condition implies that the solution must pass close to the point $(0, N_s)$ in the (S, Q) plane.

A systematic investigation of possible models must necessarily start with the simplest cases. It is for this reason that in this paper attention will be confined to homogeneous models, that is, models of uniform composition. We have to investigate the models mathematically and then having seen to what physical conditions they correspond, select those which are physically plausible.

3. In this, as in most stellar problems, the solutions of Emden's equation are of use in the most general case. We therefore proceed to a consideration of these solutions in the (S, Q) plane, where Emden's equation is obtained by putting $N=$constant (since $N=1/n+1$) in equation (3). Thus, we have a first order differential equation in S, Q. It will be found advantageous to derive the properties of the Emden equation directly in terms of h-variables, even though the equation has been examined in many other forms by various authors.

From the definition of S and Q only the region $S\geqslant 0$ and $Q\geqslant 0$ is of interest.

It is an immediate consequence of the equation that the solutions have horizontal tangents on the lines $Q=0$ and $S+N-Q=0$ and vertical tangents

on the lines $S=0$ and $1+2S-4Q=0$. Referring to Figs. a, b and c, it will be seen that in region I both $1+2S-4Q$ and $S+N-Q$ are positive. Therefore, in this region dQ/dS is positive. In region II, $(1+2S-4Q)$ is negative while $(S+N-Q)$ is positive, so that dQ/dS is negative. In region III, both $(1+2S-4Q)$ and $(S+N-Q)$ are negative, so that dQ/dS is positive. In region IV, which is in the quadrant under consideration only if $N<\frac{1}{4}$, $(1+2S-4Q)$ is positive and $(S+N-Q)$ is negative, so that dQ/dS is negative. The gradient changes sign on crossing either of the lines $1+2S-4Q=0$ or $S+N-Q=0$.

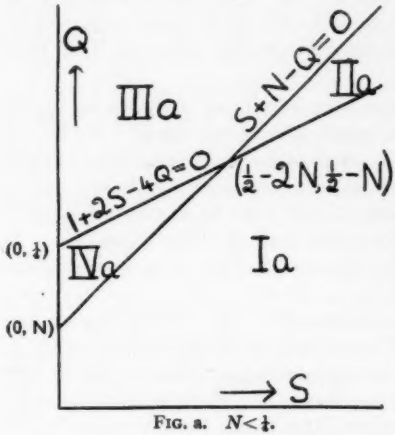


FIG. a. $N<\frac{1}{4}$.

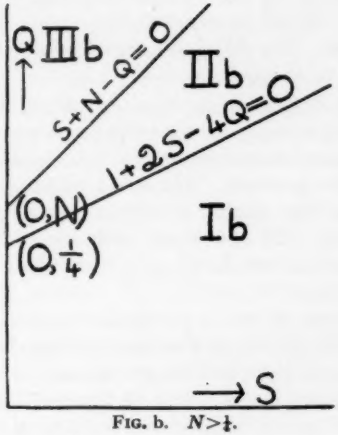


FIG. b. $N>\frac{1}{4}$.

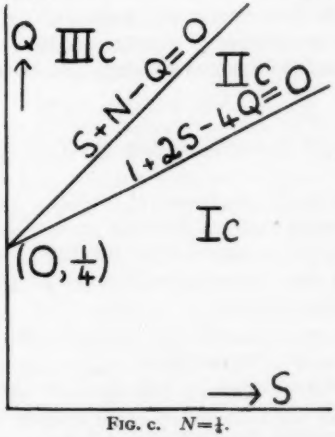


FIG. c. $N=\frac{1}{4}$.

The points $(0,0)$; $(0,N)$; $(\frac{1}{2}-2N, \frac{1}{2}-N)$; and $(-\frac{1}{2},0)$ are singular points. The fourth point is not of interest in the present connection, since it does not lie in the quadrant $S\geqslant 0$, $Q\geqslant 0$, and for the same reason the third point is of interest only if $N\leqslant\frac{1}{4}$.

(i) The point $(0,0)$. The solutions starting from this point are of the form $Q\sim AS^N$ near $(0,0)$. Such a solution starts with a vertical tangent, and for increasing S enters region I, in which dQ/dS is positive. Therefore, such a solution cannot cross the line $Q=0$ to leave region I, nor can it cross the line

$S+N-Q=0$ without first crossing the line $1+2S-4Q=0$. Thus either it must run to infinity without leaving the region, or it must cross the line $1+2S-4Q=0$ with a vertical tangent. In the first case the general solution behaves like $Q=BS^{1/2}+O(1)$ for large S. For $N>\frac{1}{6}$ there is a particular solution which behaves like $Q=S/3+O(1)$ for large S. This special solution is the Emden function $Q_E(S;N)$. It separates the solutions which do not cross the line $1+2S-4Q=0$ from those that do. For $N\leqslant\frac{1}{6}$ there are no solutions springing from $(0,0)$ that cross the line $1+2S-4Q=0$.

For $N>\frac{1}{6}$ the solutions crossing the line $1+2S-4Q=0$ enter region II in which dQ/dS is negative and cross the line $S+N-Q=0$ with a horizontal tangent. For $N\geqslant\frac{1}{4}$ the solutions run to the point $(0,N)$ while for $N<\frac{1}{4}$ they spiral to the point $(\frac{1}{2}-2N,\frac{1}{2}-N)$.

(ii) The point $(0,N)$. For $N>\frac{1}{3}$ there is an infinity of solutions springing from this point, in the neighbourhood of which Q behaves like $N+AS^{N/4N-1}$. A solution that enters region II b therefore starts with a vertical tangent and has negative gradient. There is a point of inflexion in region II b and the solution crosses into region I b with a vertical tangent and runs to $(0,0)$. A solution entering III b also starts with a vertical tangent, but has dQ/dS positive. It crosses the line $S+N-Q=0$ and then the line $1+2S-4Q=0$ and runs to the point $(0,0)$.

There is also a particular solution, analytic at $(0,N)$ which behaves like $N-NS/(3N-1)$ in the neighbourhood of the point. It also runs to $(0,0)$.

For $N=\frac{1}{3}$ there is an infinity of solutions springing from $(0,\frac{1}{3})$, in the neighbourhood of which Q behaves like $N-S\log AS$. Such a solution starts with a vertical tangent and must enter region III b (since $\log AS$ is essentially negative for small S). It then runs to the point $(0,0)$.

For $\frac{1}{4}<N<\frac{1}{3}$ there is an infinity of solutions running to the point $(0,N)$ and in the neighbourhood of this point the solutions behave like

$$Q=N+NS/(1-3N)+o(S).$$

These solutions start with a finite positive gradient, enter region III b and continue as above.

For $N=\frac{1}{4}$ the points $(0,N)$, $(0,\frac{1}{4})$ and $(\frac{1}{2}-2N,\frac{1}{2}-N)$ coincide. There is a single solution running to the point, of the form $\frac{1}{4}+S+8S^2+\ldots$ in its neighbourhood. The solution therefore starts with the line $S+\frac{1}{4}-Q=0$ as tangent, enters region III c and then crosses the line $S+\frac{1}{4}-Q=0$ with a horizontal tangent, eventually running to $(0,0)$.

For $N<\frac{1}{4}$ there is only one solution through the point $(0,N)$. In the neighbourhood of the point, the solution is $Q=N+NS/(1-3N)+o(S)$. If $\frac{1}{4}>N>\frac{1}{6}$, the solution enters region I a and spirals to the point $(\frac{1}{2}-2N,\frac{1}{2}-N)$. If $N=\frac{1}{6}$ the solution is $Q=S/3+\frac{1}{6}$ and so is the Emden function for $N=\frac{1}{6}$. If $N<\frac{1}{6}$ the solution behaves like $Q=O(S^{1/2})$ at infinity.

(iii) The point $(\frac{1}{2}-2N,\frac{1}{2}-N)$ for $N<\frac{1}{4}$. This is a spiral point. For $N=\frac{1}{4}$ this point has already been discussed.

4. We have shown in Section 2 that large increases in radius over a main sequence star of the same mass are obtained when the solution has S and $S+N-Q$ small together. Since at the surface of a star both the pressure and temperature must be negligibly small, S and Q must be very small at the surface and there is no loss of accuracy as a result of putting $S=Q=0$ at the surface. We now

consider the solutions discussed in Section 3(i) starting from (0, 0) and consider the limiting case of A infinite.

If A is sufficiently large, the solution will pass close to the point $(0, N)$. Let us consider solutions near this point, putting $S=x$, $Q=N-y$ (x and y both small). Then

$$\frac{dy}{dx} = -\frac{(N-y)(x+y)}{x(1+2x-4N+4y)}.$$

Then, if $N<\frac{1}{4}$, dy/dx is negative and dQ/dS is positive. If $N>\frac{1}{4}$, dy/dx is positive and dQ/dS is negative provided x and y are sufficiently small.

Since Q is initially less than N it is only the case of y positive that is of interest.

Thus if $N>\frac{1}{4}$, the limiting solution as A tends to infinity runs into the point $(0, N)$. But if $N<\frac{1}{4}$ the limiting solution will pass close to the point $(0, N)$ and then follow closely the analytic solution starting from $(0, N)$ and discussed in Section 3(ii). That the different form of the solutions for the two cases are to be expected can be seen from diagrams a and b.

It is therefore stars with surface value of $N<\frac{1}{4}$ and very large A values that are of interest. Stars of this type with infinite A values in fact have infinite radius and the case is a limiting one, but it is clear that stars of very large radii would be obtained by taking very large A values. To illustrate this point we consider the limiting case. Taking the model in which N is constant throughout, we obtain, as the solution with infinite radius and finite mass, the polytrope $N=\frac{1}{6}$. (The well-known Emden polytrope $n=5$.) This is the only possible solution, since for S infinite (i.e. the centre of the star) Q must behave like $S/3$. No other value of N has a solution with this behaviour running from $(0, N)$.

This solution, however, is not of great physical interest. It implies a ratio of specific heats of $\frac{6}{5}$ throughout, which is impossible for matter in the interior of stars. Also no known law of radiative transfer would give rise to a constant value of N.

5. In this section it is shown that even when N is not constant, the only possible mathematical solutions for homogeneous red giants are ruled out on the grounds that the opacity laws which would be necessary mathematically do not correspond to the physical laws which are believed to be correct. It is believed that at high temperatures (10^6 to 10^7 degrees) the opacity is mainly due to the photoelectric effect and varies like $\rho/T^{3\cdot5}$, while at lower temperatures some correction tending to reduce the power of ρ and raise the power of T in this expression, is required. (The variation of the guillotine factor.) In order to consider a definite mathematical problem, the existence of homogeneous stars of *infinite* radius is first considered and disproved, but it is shown finally that the solutions for stars of large but non-infinite radii lie so extremely close to the infinite solution that the disproof is applicable to them also.

If N is not constant, the solutions of equation (3) with N constant are still of interest in giving some idea of how the curves run in the (S, Q) plane over a small range of S. In particular these solutions give the form of the curves for small S.

We are considering only models of uniform composition and so by definition Q and N and therefore dQ/dS must be continuous. Thus models with electron opacity or Kramers opacity throughout the radiative region cannot give extensions of radii, since the only solutions which fit on smoothly to a convective core are those described in Paper A which correspond to the standard model

or main sequence stars. Nor can they correspond to a star with mixed photoelectric and electron opacities, since such models are discussed elsewhere * and are shown to have properties lying between the two extreme cases.

In the inner part of the radiative region it is known that the only possible opacity laws are those of photoelectric effect or electron scattering, or some combination of them. Therefore on physical grounds the only remaining possibility of obtaining large radii for homogeneous stars lies in the consideration of stars in which the atmosphere has a different opacity law in the exterior from that in the interior. The most likely explanation of this fact is that in the inner part of the atmosphere the photoelectric law should apply, while in the outer part where the temperature is lower a guillotine factor correction should be included.

The photoelectric opacity is dependent on $\rho/T^{3\cdot5}$ and leads to an energy transport equation in h-variables †

$$\frac{S}{N}\frac{dN}{dS} = \frac{S+2-8\cdot5N}{1+2S-4Q}. \tag{4}$$

When the guillotine factor is introduced, the dependence of the opacity on the density and temperature can be expressed as

$$\frac{\rho^{\alpha\beta-1}}{T^{\beta(1-\alpha)-3}},$$

where α, β depend on the composition and state of the material. We shall not consider in detail the most probable values of α and β on physical grounds, except to indicate their extreme values, and will concentrate on which values give rise to large radii.

With the opacity depending on ρ and T as above, the energy transport equation is

$$\frac{S}{N}\frac{dN}{dS} = \frac{S+\beta(\alpha-N)}{1+2S-4Q}. \tag{5}$$

If we consider solutions of equations (3) and (5) ‡ in the (S, Q) plane, there is an infinity of solutions starting from the origin and Q behaves like AS^{α} for small S, as in the constant N case. The surface value of N is α.

The limiting solution with A infinite will run along $S=0$ as far as $Q=\alpha$. At the point $(0, \alpha)$, N is still equal to α, since $dN/dQ=0$ when $N=\alpha$ and $S=0$. Thus the solution starts similarly to that discussed at the beginning of Section 4. α can be assumed to be less than $\frac{1}{4}$, since for pure Kramers opacity it is $\frac{4}{17}$ and the effect of the guillotine factor is to reduce it.

As is shown in Section 3, when $N<\frac{1}{4}$ there is only one solution of the Emden equation through the point $(0, N)$. When N is not constant, but satisfies equation (5), there is again only one solution through $(0, \alpha)$, although the term of order S in Q has no longer the same coefficient. We can obtain expansions of Q and N in powers of S which are valid for small S, and then continue to integrate the solution in the usual way.

Let us consider a complete solution in which α and β for the outer part of the atmosphere are given. To give some idea of the possible values of α, it should be remarked that all physical theories indicate that the opacity must be an increasing function of ρ and a decreasing function of T. Hence if α is small,

* C. M. Bondi, *M.N.*, **109**, 571, 1949.

† Paper A, equation (11).

‡ These equations are considered more fully in the appendix.

β must be large, since the opacity varies like $\rho^{\alpha\beta-1}$. The opacity also varies like $1/T^{\beta(1-\alpha)-3}$, so that the opacity is in fact a function of $\rho/T^{\{[\beta(1-\alpha)-3]/[\alpha\beta-1]\}}$. It seems likely that the power of T in this expression should not be too great (in the ordinary Kramers opacity law it is 3·5) and so we can assume that for any physically possible law $\alpha \geqslant 0{\cdot}1$. As an example, if the opacity varies like $\rho^{3/4}/T^4$, then $\alpha = 0{\cdot}2$, while $\beta = 8{\cdot}75$.

We shall consider separately the three parts of the solution in the (S, Q) plane. Part (*c*) is the Emden solution for the convective core with $n = 1{\cdot}5$ (i.e. $N = 0{\cdot}4$), part (*b*) is one of the single infinity of non-intersecting solutions with Kramers opacity which fit smoothly on to the Emden solution, and part (*a*) is the guillotine factor solution. Only one of solutions (*b*) will fit smoothly on to the solution (*a*) for the assumed α and β. One particular solution (*b*) is the standard Kramers solution running to the point (0, 0). This will be referred to as the standard solution. The others are automatically divided into two classes, those lying above the standard and those lying below it. Those lying above it eventually spiral to the point $S = \frac{1}{2}$, $Q = \frac{1}{2}$, where $N = 0$. N decreases all the time along these curves. Those below it have a vertical tangent where N goes to infinity and then run with S increasing and are asymptotically equal to $Q = 0$. Solutions (*b*) can be characterized by the point at which they leave the Emden solution (*c*).

In fact, as will be shown later, only those solutions (*b*) which leave solution (*c*) above the standard (i.e. with $S > 2{\cdot}12$) are of interest and we shall consider only these. Let us consider two such solutions, one leaving the Emden solution at (S_1, Q_1) and the other at (S_2, Q_2), where $S_1 > S_2$. Now since at the point of fit of solutions (*b*) and (*c*), N is continuous, the two solutions (*b*) and (*c*) must fit together smoothly. Since the Emden solution for $S > 2{\cdot}12$ is almost linear, Q varying like $S/3$, this implies that provided (S_1, Q_1) and (S_2, Q_2) are not very far apart, the solution (*b*) leaving at (S_1, Q_1) must pass close to and slightly above the point (S_2, Q_2). Thus, in particular, the solution (*b*) leaving the Emden at (S_1, Q_1) passes near to the point $(2{\cdot}12, 0{\cdot}82)$ where the standard solution leaves the Emden.

It can be seen from the Kramers equations that the rate of change with S of $NQ^{8{\cdot}5}$ is a function of S and Q only. As the solution (*b*) proceeds from the Emden in the direction of S decreasing, N remains close to the standard N and then diverges rapidly from the standard values, while Q only does so slowly and then only in consequence of the divergence of N. Thus until Q has diverged from the standard, $NQ^{8{\cdot}5}$ is very close to the standard value.

Let us consider now solution (*a*). It is impossible for N to drop below the value α which it takes at the surface. For, if we consider equation (5), we have that dS and $(1 + 2S - 4Q)$ have the same sign (from equation (3)), S and N are positive and so dN and $\alpha - N + S/\beta$ have the same sign. Hence, when N is increasing, $N < \alpha + S/\beta$, and when N is decreasing, $N > S/\beta + \alpha$. Now again, since S is positive, this gives the result that on any solution (*a*), N is always at least as great as α.

At the point of fit of solutions (*a*) and (*b*), N is to be continuous. Therefore, at this point the value of N on (*b*) must be at least equal to α and we have taken for the minimum possible of α the value 0·1.

Therefore, from the previous argument, since $N_{ab}Q_{ab}^{8{\cdot}5}$ (where suffix *ab* refers to the point of fit of solutions *a* and *b*) is very close to the standard value on

solution (*b*), at the point of fit, Q_{ab} must be close to the standard value for the same value of S and in fact differs from it by a factor at most $(10N_{ab})^{1/8\cdot5}$.

Also, on solution (*a*) by an argument similar to that for N, Q never drops below α. Hence Q_{ab} must be greater than or equal to α and hence by our assumption for α cannot be below 0·1.

It now becomes of interest to examine the standard solution at $Q=0\cdot1$. For this is the lowest possible value of Q at which a fit on to solution (*a*) can take place, and as we have shown, the actual solution at this point cannot be far from the standard, differing from it by a factor $(10N_{ab})^{1/8\cdot5}$ at most.

At $Q=0\cdot1$ the standard solution (*b*) has S approximately 0·03 and T/T_c is equal to approximately 0·2. T/T_c is obtained from the quadrature

$$\log T_c/T=\int_S^\infty \frac{N\,dS}{S(1+2S-4Q)}.$$

Over a short range of S the value of N in (*b*) will be less than that in the standard solution, while $1+2S-4Q$ is very close to the standard. (The standard solution is nowhere near the line $1+2S-4Q=0$.) Hence the value of the integral is slightly less than the value of the corresponding integral for the standard solution. Hence T/T_c is slightly greater than the standard value. It is found below that it is necessary only to use the fact that T/T_c at $Q=0\cdot1$ is not less than 0·2, so that no further quantitative investigation of the value of T/T_c is required.

It is known that the central temperatures of stars cannot differ very much from each other. Thus taking the central temperature as that of the Sun, i.e. 2×10^7 degrees, we have shown that if a star of uniform composition can have a very large radius, the guillotine factor must apply at a point in the star where the temperature is at least equal to 4×10^6 degrees. On physical grounds this can be ruled out, so that this would seem to be a disproof of the existence of stars of very large radii with uniform composition, at least so far as Kramers opacity interiors are concerned.

A few general considerations show that stars with electron opacity interiors lead to similar contradictions. For if one follows electron opacity solutions outwards, a temperature must eventually be reached at which Kramers opacity is important and this must of course occur before the guillotine factor becomes dominant. Now in an earlier paper* some models with mixed opacity have been integrated and it is shown that the mixed solutions have properties lying in between those of the pure Kramers and pure electron opacity solutions. On the pure electron opacity solution Q is 0·1 at $S=0\cdot04$ and this corresponds to a value of T/T_c of 0·196. Hence the value of T/T_c at the inner boundary of the guillotine factor solution is only very slightly affected by the introduction of a mixed opacity law and the same considerations apply as before.

Integrations show that an order of magnitude solution can be given for a large but non-infinite star. If (S, Q) denotes a point on such a solution, (S, Q^*) the point on a standard (main sequence) solution for the same S value and (S, Q') the point on the solution starting from $(0, N_s)$ for the same value of S, the extension of the large radius solution over the standard solution is of the order of

$$\frac{Q'-Q^*}{Q'-Q},$$

* C. M. Bondi, *M.N.*, **109**, 571, 1949.

where S is taken at the point where the large radius solution has its inflexion. This shows how very close a solution giving, say, an extension of 50 to 100 must be to the infinite radius solution. Since it is impossible, by a considerable margin, to fit the infinite (*a*) solution on to a (*b*) solution, this shows that it is also impossible to fit a large radius (*a*) solution on to a (*b*) solution.

APPENDIX

Solutions (*a*) have not been considered elsewhere and so form the subject of a short discussion.

TABLES

$\alpha=0{\cdot}1$, $\beta=0{\cdot}1$

S	Q	N
0	0·1	0·1
0·05	0·1082	0·1080
0·10	0·1162	0·1159
0·20	0·1316	0·1311
0·60	0·1877	0·1862
1·00	0·2373	0·2348

$\alpha=0{\cdot}1$, $\beta=1$

S	Q	N
0	0·1	0·1
0·05	0·1080	0·1070
0·10	0·1158	0·1138
0·20	0·1309	0·1268
0·60	0·1852	0·1727
1·00	0·2328	0·2122

$\alpha=0{\cdot}1$, $\beta=10$

S	Q	N
0	0·1	0·1
0·05	0·1075	0·1030
0·10	0·1147	0·1060
0·20	0·1284	0·1115
0·60	0·1765	0·1306
1·00	0·2170	0·1464

$\alpha=0{\cdot}14$, $\beta=2$

S	Q	N
0	0·14	0·14
0·05	0·1544	0·1495
0·10	0·1688	0·1590
0·20	0·1979	0·1774
0·60	0·3160	0·2475
1·00	0·4369	0·3139
1·60	0·6237	0·4106

$\alpha=0{\cdot}15$, $\beta=1$

S	Q	N
0	0·15	0·15
0·05	0·1676	0·1634
0·10	0·1858	0·1774
0·20	0·2247	0·2062
0·40	0·3143	0·2693
0·70	0·5443	0·4090

$\alpha=0{\cdot}15$, $\beta=5$

S	Q	N
0	0·15	0·15
0·05	0·1654	0·1564
0·10	0·1809	0·1627
0·20	0·2121	0·1748
0·40	0·2751	0·1978
1·00	0·4688	0·2614
2·00	0·8117	0·3607
2·60	1·0330	0·4190

In the neighbourhood of the point $S=0$, $Q=\alpha$, Q and N can be expressed as a series of powers of S. To second order,

$$Q=\alpha+a_1S+a_2S^2, \quad N=\alpha+b_1S+b_2S^2,$$

where

$$a_1 = \frac{\alpha(1-3\alpha+\alpha\beta)}{(1-3\alpha)(1-4\alpha+\alpha\beta)}, \qquad b_1 = \frac{\alpha}{1-4\alpha+\alpha\beta},$$

$$b_2 = \frac{\alpha(10\alpha^2\beta - 24\alpha^2 - 2\alpha\beta + 10\alpha - 1)}{(1-3\alpha)(1-4\alpha+\alpha\beta)^2(2-8\alpha+\alpha\beta)},$$

$$a_2 = \frac{\alpha b_2 + a_1(3a_1 + b_1 - 1)}{2-7\alpha}.$$

Using these expressions and then integrating step by step, a number of integrations were obtained for $\alpha = 0{\cdot}1$, $0{\cdot}14$ and $0{\cdot}15$. Those for $\alpha = 0{\cdot}1$ were continued as far as $S = 1$. Those for $\alpha = 0{\cdot}14$ and $0{\cdot}15$ were continued until $N = 0{\cdot}4$.

It is of interest to note that if the guillotine opacity solution were to fit directly on to the core (that is with no intermediate solution (*b*)), the condition that $\beta \geqslant 0$ gives at once the restriction on α, $0{\cdot}131 \leqslant \alpha \leqslant \frac{1}{6}$.

Newnham College,
Cambridge :
1950 *April* 29.

MODELS FOR RED GIANT STARS

II. Models with a Chemical Inhomogeneity and Opacity due to Photoelectric Effect

C. M. Bondi and H. Bondi

(Received 1950 May 4)

Summary

This paper presents a complete survey of the range of models that can be constructed using the opacity law $\kappa \sim \rho^{0.75} T^{-3.5}$ and possessing a single chemical discontinuity. It is shown that for physically possible changes of molecular weight, large and even theoretically infinite radii are possible with only a moderate proportion of the whole mass in the outer envelope. The whole system of models is surveyed and mapped diagrammatically in homology-invariant variables, thereby showing the ranges of the various parameters important to the red giant problem. The solutions recently given by Li Hen and Schwarzschild, while numerically correct, are shown to be unrepresentatively distributed and confined to the less interesting part of the diagram. The outer regions of stars of large radii are further considered, and it is concluded that breakdown of the opacity law is likely to set a physical upper limit to the actual radii of red giant stars.

---

1. It was shown by Hoyle and Lyttleton in 1942 (**1**) that the large radii of red giant stars were likely to be due to the outer region of the star having lower molecular weight than the inner region. The mathematical difficulties standing in the way of a complete survey of the consequences of this hypothesis are very considerable, and these authors, in their initial paper, were content to show that in certain circumstances very large radii could result. In a more recent paper (**2**) they have carried the investigation a great deal further, considering the cases both of photoelectric effect and of electron opacity. Another recent paper by Li Hen and Schwarzschild (**3**) considers the case of photoelectric effect (Kramers' opacity), slightly modified to allow for the variation of the guillotine factor, but claims to establish that only moderate extensions of radius can occur. It is the purpose of the present paper to carry out a complete survey of the problem by means of the technique developed in a previous paper (**4**) and first applied to the red giant problem in another paper (**5**). It will be shown that, adopting the same law of opacity as Li Hen and Schwarzschild, a sufficiently large, but physically entirely possible, change in composition can lead to *an infinite extension of radius* and that even for smaller changes in composition large radii will occur. The contrary conclusion reached by Li Hen and Schwarzschild is shown to be incorrect; it is also shown that these authors have been led to it as a result of inferences drawn from the unrepresentative set of integrations performed by them. The mathematical agreement between their models and our survey is satisfactory, but the selection of their models was unfortunately restricted to a comparatively uninteresting set of cases. In our view only some such method as is used in the present paper can lead to a complete survey of the problem and so establish what a truly representative range of models is.

2. It was shown in a previous paper (**5**) that no physically plausible variation of the law of opacity could lead to a ***homogeneous*** star of large radius with nuclear generation of energy. The next simplest model is a star with a single discontinuity of composition, the surface of discontinuity being a sphere of radius r_2 concentric with the star. This is the model proposed by Hoyle and Lyttleton (**1**). Physically a star closely represented by this model is likely to arise when a star that has converted a substantial fraction of its hydrogen into helium then accretes appreciable amounts of interstellar matter mainly composed of hydrogen. The proportion by mass of elements other than hydrogen and helium is unaffected by the thermonuclear processes taking place near the centre and may therefore be taken to be the same in the interior as in the envelope; this proportion will, as usual, be denoted by Z and is probably of the order of 1 per cent. The helium abundance by mass Y in interstellar material is negligibly small, but in the stellar interior may have any value up to the extreme one of $1-Z$ if all the hydrogen X has been converted into helium.

In this paper we shall neglect radiation pressure. The opacity law considered will be Kramers' law modified to take some account of the variation of guillotine factor as adopted by Li Hen and Schwarzschild.

Hence

$$\kappa=\kappa_0\rho^{0\cdot75}T^{-3\cdot5}, \qquad \kappa_0=\kappa_{00}Z(1+X)^{0\cdot75}, \qquad \kappa_{00}\doteqdot 10^{25}\text{ c.g.s. units.} \tag{1}$$

The energy transport equation in physical (p-) variables is therefore *

$$\frac{dT}{dr}=-\frac{3\kappa_0 L}{16\pi ac}\frac{\rho^{1\cdot75}}{r^2T^{6\cdot5}}. \tag{2}$$

At the interface $r=r_2$ the variables P, T, M and L, and hence $S=4\pi Pr^4/GM^2$, must all be continuous. It follows that ρ/μ and $Z^{-1}\mu^{-1\cdot75}(1+X)^{-0\cdot75}\,dT/dr$ are continuous. The values of the discontinuous variables ρ, $Q=Pr/GM\rho$ and $N=1-d\log\rho/d\log P$ just outside and just inside the interface will be denoted by the suffixes e, i respectively. Similarly the values of X, Y, Z and μ in the exterior and interior will be distinguished by the suffixes e, i respectively.

Then

$$\frac{\rho_{2e}}{\rho_{2i}}=\frac{\mu_e}{\mu_i}, \qquad \frac{Q_{2e}}{Q_{2i}}=\frac{\mu_i}{\mu_e}, \qquad \frac{N_{2e}}{N_{2i}}=\frac{Z_e}{Z_i}\left[\frac{\mu_e(1+X_e)}{\mu_i(1+X_i)}\right]^{0\cdot75}. \tag{3}$$

These conditions are equivalent to those used by Li Hen and Schwarzschild. The last condition applies only if the equilibrium is radiative on both sides of the interface, as is usually the case. Also, in common with those authors, it will be assumed that the energy generation is confined to the central convective core which must of necessity exist.

The star may conveniently be divided into three zones as follows, beginning at the surface $r=r_s$:—

(*a*) $r_s\geqslant r>r_2$: an envelope of composition (X_e, Y_e, Z_e; μ_e). This envelope will be found to be in radiative equilibrium. (In a small and unimportant class of models a minor part of it may be convective.)

(*b*) $r_2\geqslant r>r_1$: an intermediate zone of composition (X_i, Y_i, Z_i; μ_i) in radiative equilibrium.

(*c*) $r_1\geqslant r\geqslant 0$: an inner convective core having the same composition as zone (*b*).

* The notation and the fundamental equations are taken from (**4**).

Since zone (*c*) always exists, the range of cases arising depends on the relative extents of zones (*a*) and (*b*). In the usual case the boundary between (*a*) and (*b*) will be between two zones in radiative equilibrium. If, as is to be expected, $\mu_e<\mu_i$ the interface is stable, the lighter material overlying the heavier one. The discontinuity will thus be sharp. If zone (*a*) becomes vanishingly small the star is homogeneous, and the solution in this case will be referred to as the *standard solution.*

3. The solution in the convective core (*c*) is well known, being given by the Emden function

$$Q=Q_E(S;\ 0{\cdot}4), \qquad N=0{\cdot}4. \tag{4}$$

The solution in zone (*b*) must satisfy the hydrostatic equation

$$\frac{dQ}{dS}=\frac{Q}{S}\frac{S+N-Q}{1+2S-4Q}, \tag{5}$$

and the energy transport equation which, with the opacity law here assumed (*cf.* (**4**), Part I, equation (27)), takes the form

$$\frac{dN}{dS}=\frac{N}{S}\frac{S+1{\cdot}75-8{\cdot}25N}{1+2S-4Q}. \tag{6}$$

Only a one-parameter family of the solutions of the second order system (5, 6) is of interest in zone (*b*). For, since at $S=S_1$ no change of composition occurs, all the variables must be continuous, and hence only those solutions of (5) and (6) that join smoothly on to (4) at $S=S_1$ need be considered, and these solutions need only be investigated for $r\geqslant r_1$, which in fact implies $S\leqslant S_1$. The value of S_1 may be taken as the parameter specifying the members of the family. One member of the family with parameter $S_1=S^*$ is the standard solution

$$Q=Q^*(S), \qquad N=N^*(S) \tag{7}$$

which runs to the surface $S=Q=0$, $N=\frac{7}{33}$.

It is a crucial property of this family of solutions that if at any point $S=S_a$ the difference $N-N^*$ is of the same order of magnitude as N^*, then at that point and for all $S>S_a$, Q is very close to Q^*. Hence $Q=Q^*(S)$ gives a good approximation to all members of the family whose values of $N(S)$ do not differ by a large factor from $N^*(S)$. An outline of a proof of this vital property is as follows:—

It can be seen by an analytic consideration of small variations of the solutions of (5) and (6) that, in the neighbourhood of the standard solution, as S decreases, (6) is highly unstable, but (5) is moderately stable. Hence if a member of the family lying very close to the standard solution near $S=S_1^*$ is followed towards lower values of S then, once N begins to deviate appreciably from N^* it will rapidly depart substantially (such behaviour may be termed "slicing off"). On the other hand, owing to the stability of (5), Q will stay close to Q^* until the departure of N from N^* so alters the right-hand side of (5) that dQ/dS deviates markedly from its standard values. It is only when the effects of the differences between the actual and the standard values of dQ/dS have accumulated over an appreciable range of S that Q will depart significantly from its standard values. Hence by the time N differs from N^* by any number smaller than N^*, Q will still be close to Q^*.

A higher approximation is easily found. Consider

$$\frac{d}{dS}\log(NQ^{8\cdot25}) = \frac{9\cdot25S - 8\cdot25Q + 1\cdot75}{S(1+2S-4Q)}.$$

Since the right-hand side of this equation does not contain N, the values of $NQ^{8\cdot25}$ will only depart appreciably from its standard values when Q has deviated markedly for a significant range of S. Therefore $NQ^{8\cdot25}$ will retain its standard values even better than Q and much better than N. The equation

$$NQ^{8\cdot25} = N^*Q^{*8\cdot25} \tag{8}$$

can be shown, by rough computation of one or two numerical examples, to hold within a fraction of one per cent as long as $(N-N^*)$ is less than N^*. This is all the accuracy likely to be required. Hence no integrations (other than for the standard model) are required in the intermediate zone (*b*). Formula (8) takes, in our work, the place of all the integrations of Li Hen and Schwarzschild, but covers all cases instead of only six. Formula (8) also demonstrates the smallness of the deviation of Q from Q^*, a fact which Hoyle and Lyttleton (**1**), (**2**) have implicitly recognized and consistently used in their work.

4. The solutions in the envelope (*a*) also form a one-parameter family of solutions of (5) and (6) in the (S, Q) plane, since they must all run to the surface $S=0$, $Q=0$, $N=\frac{7}{33}$. Near the surface these solutions are given approximately by

$$Q = AS^{7/33} + O(S^{14/33}), \qquad N = \tfrac{7}{33} + O(S) \text{ as } S \to 0, \tag{9}$$

where A is a constant, which we may make use of to label the members of the family. This family and the one mentioned in the preceding section have only one member in common, namely the standard solution (7) for which $A = A^*$. The solutions of interest for the present problem are only those for which $A \geqslant A^*$.

The general behaviour of these solutions is perhaps best understood by considering first a non-member of the family, the special solution F discussed in a previous paper (**5**). This is the unique solution that starts at $S=0$, $Q=\frac{7}{33}$ with $N=\frac{7}{33}$. In the neighbourhood of that point

$$Q = \frac{7}{33} + \frac{651S}{1004} + \ldots, \qquad N = \frac{7}{33} + \frac{28S}{251} + \ldots. \tag{10}$$

The solution then spirals to the point $S=\frac{5}{82}$, $Q=\frac{23}{82}$ with $N=\frac{9}{41}$. All members of the family with greater A than a certain value $\bar{A}(\bar{A}<A^*)$ spiral to this same point. If $A \leqslant \bar{A}$ the solution runs to infinity. No two members of the family intersect and successive turns of the spiral are separated by the curve F. The curves fill the entire quadrant of the (S, Q) plane $S \geqslant 0$, $Q \geqslant 0$ and one and only one member of the family passes through any given point of the plane other than the points of F. The larger A the more steeply the curve rises near the origin. Since the curves, however, cross neither the Q-axis nor the spiral F they all turn sharply right below the spiral, in whose vicinity they afterwards remain (*cf.* Fig. 1). Hence a specific value of N corresponds to each point of the plane, namely the value N attains at that point when reached from the origin along the unique member of the family passing through the point. Curves of constant N can be drawn on a diagram representing all the envelope solutions (Fig. 1). It will be seen that the variation of N is slow and smooth.

As a solution is followed into the star from the origin, $dS/(1+2S-4Q)$ is always positive, and so are N and S. Hence N will be increasing or decreasing according as the sign of $(S+1{\cdot}75-8{\cdot}25N)$ is positive or negative. It follows that N cannot attain any value $N_a > \frac{7}{33}$ unless $S > 8{\cdot}25N_a - 1{\cdot}75$. Hence convective instability* cannot occur in the envelope for $S < 1{\cdot}55$, nor, with the discontinuity conditions for N to be discussed in Section 4, can such instability occur on the outer fringe of the intermediate zone (b) for S less than about 0·9. It will be shown later that large radii occur only for fairly small values of S_2 (about 0·1), so that convective instability need not be considered any further in this paper. For these low values of S, the value of N hardly differs from its surface value and this justifies the assumption of its constancy made by Hoyle and Lyttleton (**2**).

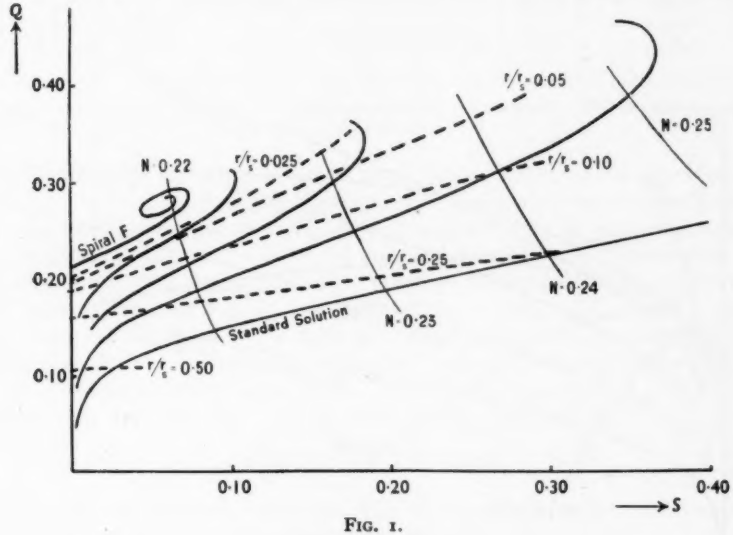


FIG. 1.

Solutions.
Curves of constant N.
Curves of constant r/r_s.

Just as in the case of N, so also in the cases of (*cf.* equation (12), Part II of (**4**))

$$\frac{M}{M_s} = \exp\left\{-\int_0^S \frac{dS}{1+2S-4Q}\right\}, \qquad \frac{r}{r_s} = \exp\left\{-\int_0^Q \frac{dQ}{S+N-Q}\right\}, \qquad (11)$$

a specific value of each ratio may be associated with every point of the plane by carrying out the integration from the origin to the point along the unique member of the family passing through the point. Hence curves of constant M/M_s and curves of constant r/r_s may be drawn on the diagram representing the envelope solutions (Figs. 1, 2). The curves of constant M/M_s start on the convex side of the spiral F and run to the S-axis, reaching it at the point for which $(1+2S)^{-1/2} = M/M_s$ (since $Q=0$ is a member of the family). On the section of

* Convective instability cannot occur as long as the polytropic index n exceeds 1·5. Since, by equation (3) of (**5**), Part I, $N = 1/(n+1)$, it cannot occur for $N < 0{\cdot}4$.

the Q-axis between the origin and $Q=\frac{7}{33}$ we have $M/M_s=1$, while for $S=0$, $Q>\frac{7}{33}$, and on the concave side of the spiral F the value of M/M_s vanishes.

The curves $r/r_s=$const. all start on $S=0$ with $Q=7(1-r/r_s)/33$ and run to infinity. On $Q=0$ we have $r/r_s=1$, on the spiral F we have $r/r_s=0$. For low values of r/r_s the curves of constant r/r_s spiral in a certain way between the turns of F and emerge again, running out to infinity. Note that r/r_s is a continuous function of S and Q, but that M/M_s is discontinuous on F.

In the red giant problem we are chiefly interested in envelopes occupying a large part of the radius of the star, but containing only a moderate fraction of its mass. It will be seen that such envelopes must terminate with values of S and Q corresponding to points near the convex side of F and close to its beginning, i.e. close under the early part of the spiral F. This is just the region found to be important to the red giant problem in a previous paper (**5**).

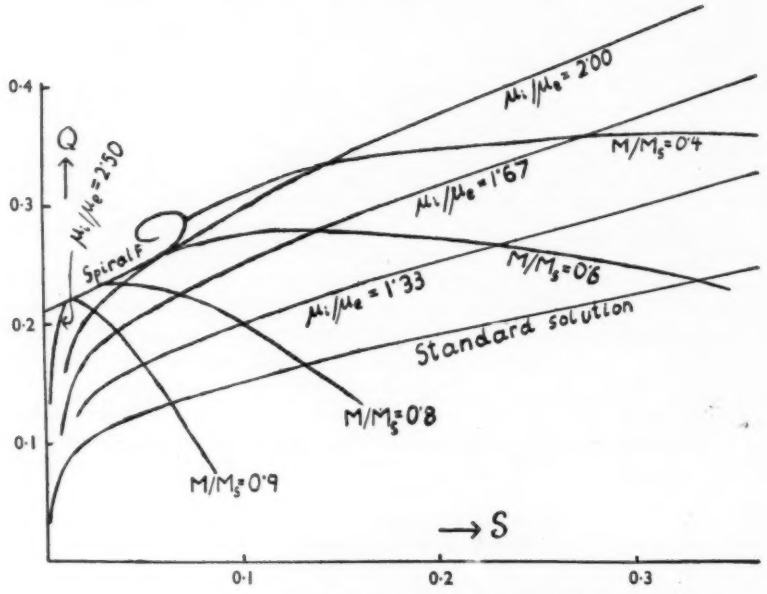


FIG. 2.—*Fitting curves. Curves of constant M/M_s.*
The fitting curves are drawn for $Y_e=0$, and apply for Z between 0 and 0·1.

5. In order to construct complete stellar models, the envelope curves have to be fitted on to the solutions for the intermediate zone. As shown above S, μQ, $Z\mu^{0\cdot75}(1+X)^{0\cdot75}/N$ are each continuous at the interface. In order to arrive at definite values for the magnitudes of the jumps of Q and N, the ratios Q_{2e}/Q_{2i} and N_{2e}/N_{2i} are tabulated below for some cases of interest. As stated before, attention has been concentrated on cases where the abundance of elements, other than hydrogen and helium, is the same in the interior and in the envelope, so that $Z_e=Z_i$. It has also been assumed that there is no helium in the envelope, so that $Y_e=0$. This will correspond to the hypothesis that the envelope consists of accreted interstellar hydrogen. According to current views on stellar structure, Z

is almost certainly less than 10 per cent, and in order to bracket likely cases the tabulations have therefore been carried out for $Z=0{\cdot}10$ and for the limiting case $Z=0$. The jumps are found to depend so insensitively on Z that any intermediate case can be interpolated with ease.

TABLE I

	$Z_e=Z_i=0$				$Z_e=Z_i=0{\cdot}10$			
Y_i	$\frac{Q_{2e}}{Q_{2i}}=\frac{\mu_i}{\mu_e}$	$\frac{N_{2e}}{N_{2i}}$	$\frac{Q_{2e}}{Q_{2i}}\left(\frac{N_{2e}}{N_{2i}}\right)^{1/0\cdot25}=\frac{Q_{2e}}{Q_2^*}$	Y_i	$\frac{Q_{2e}}{Q_{2i}}=\frac{\mu_i}{\mu_e}$	$\frac{N_{2e}}{N_{2i}}$	$\frac{Q_{2e}}{Q_{2i}}\left(\frac{N_{2e}}{N_{2i}}\right)^{1/0\cdot25}=\frac{Q_{2e}}{Q_2^*}$
0	1	1	1	0	1	1	1
0·25	1·186	0·972	1·182	0·25	1·203	0·968	1·198
0·50	1·455	0·937	1·444	0·50	1·510	0·922	1·495
0·75	1·882	0·888	1·856	0·75	2·028	0·848	1·988
1·00	2·667	0·813	2·601	0·90	2·552	0·802	2·484

The last row in each table gives the limiting case in which X_i vanishes. The last column gives the jump to the standard solution in second approximation.

In the envelope the solution will be given by one of the curves described in Section 4. At $S=S_2$ both Q and N jump in accordance with the tabulated ratios. The intermediate solution to which the curve jumps will, as explained above, have its Q value close to the Q of the standard solution. Its N value will be somewhat larger than the N value of the standard solution, mainly owing to the jump in N. The ratio N_2^*/N_{2i} will, however, be a little smaller than N_{2e}/N_{2i}, since, owing to the slope of the lines $N=$ const. in the envelope diagram, N_{2e} slightly exceeds N^*. This effect is very small unless S_2 is fairly large, a case of little importance, as will be shown later. For moderate S_2, then, N_{2i}/N_2^* will be between unity and, say, 1·35. By equation (7) this implies that

$$1 \geqslant \frac{Q_{2i}}{Q_2^*} \geqslant (1{\cdot}35)^{-1/0\cdot25} \doteqdot 0{\cdot}965. \qquad (12)$$

This shows how extremely small the error is when Q_{2i} is taken to lie on the standard curve, and how easily account can be taken of this error.

Now consider a specified change in composition. If the point (S_2, Q_{2i}) were on the standard curve, then the locus of points (S_2, Q_{2e}) would lie on a curve that would simply be the standard curve scaled up by a factor μ_i/μ_e. The small correction due to the difference between Q_{2i} and Q_2^* can then easily be applied if such accuracy is desired. In this way the locus of points (S_2, Q_{2e}) may be obtained, and will be referred to as the "fitting curve".

If now the fitting curve is drawn in on the S–Q diagram for the envelope diagram displaying the curves M/M_s ((Fig. 2), then the curve representing the stellar model having an assigned fraction of the stellar mass in the envelope is easily selected. The intersection of the fitting curve and the desired $M_2/M_s=$ const. curve gives the point (S_2, Q_{2e}). The values of A, N_{2e} and of r_2/r_s may immediately be read off on the diagrams, and the curve representing the outer part of the star is found. The jump conditions give Q_{2i} and N_{2i}. The intermediate curve can then be constructed, if such accuracy is required, but for most purposes it will be sufficient to follow the standard curve. In particular the values of M/M_s and r/r_s on the standard curve at the point (S_2, Q_2^*) are of

interest. The ratio $(M_2^*/M_s^*)/(M_2/M_s) = M_s/M_s^*$ is the ratio of the mass of the model under consideration to the mass of a homogeneous star of the same central temperature, central density and central composition. Similarly $(r_2^*/r_s^*)/(r_2/r_s) = r_s/r_s^*$ is the ratio of the corresponding radii. The complete run of the p-variables can be found in the usual way by quadratures.

6. The nature of the stellar models derived in this way depends critically on whether the fitting curve intersects the spiral F or not. It is clear that with a sufficiently large ratio Q_{2e}/Q_{2i} the fitting curve will intersect the spiral F, but the important question is whether such large ratios are physically possible. The minimum necessary value of Q_{2e}/Q_{2i}, by what has been said, will be very close to the minimum ratio of the Q values of the spiral F and the standard solution (for the same value of S). Tables of the spiral and the standard solution are given at the end of the paper, Table II (i) and (ii). It is found that, with the opacity law here assumed, the minimum ratio is just under 2. Applying the required small correction for the difference between Q^* and Q_{2i}, the minimum ratio of $Q_{2e}/Q_{2i} = \mu_i/\mu_e$ for the fitting curve to intersect the spiral F is 2·00.* It will be seen that such a ratio is entirely possible, though it corresponds to a rather high helium abundance (about 80 per cent) in the central regions of the star. Consider then the fitting curve for a change in composition corresponding to a ratio $\mu_i/\mu_e > 2{\cdot}00$. If the envelope contains only a small fraction of the mass of the star, the jump will occur for rather small values of S and Q. Owing to the small slope of the curves $r/r_s = \text{const.}$ the value of r_2/r_s will be appreciably smaller than r_2^*/r_s^*, and so the radius of the star will be appreciably greater than that of the standard model with the same T_c, ρ_c, μ_c. The steepness of the $M/M_s = \text{const.}$ curves in this part of the diagram ensures that the increase in mass is only small. If models with higher fractions of the masses in the envelope are considered, the jump point (S_2, Q_{2e}) will move closer and closer to the point where the fitting curve intersects the spiral F, and will reach it for a definite finite value of M/M_s which will be denoted by ν. But the values of r_2/r_s decrease to zero as the spiral is approached. *Hence, although the mass in the envelope is finite, the radial extension of the envelope tends to infinity.* Accordingly, arbitrarily large stellar models can be constructed by allowing the fraction of mass in the envelope to approach the critical value ν. The limiting case is a star of infinite radius but finite mass. The question of the luminosity of the star and the possible physical limitation of the mathematically infinite extent will be examined later.

Up to this point M/M_s has been a monotonically decreasing function of the distance from the origin as the point (S_2, Q_{2e}) moved along the fitting curve. But as the fitting curve crosses the spiral F the ratio M/M_s drops discontinuously to zero and then begins to rise again. Its detailed behaviour depends on the number of turns of the spiral intersected by the fitting curve, but it seems highly probable (though this point has not been rigorously established) that M/M_s attains nowhere on the fitting curve values in a range the upper bound of which is ν, and it is certain that such values are not attained in the neighbourhood of the point where the fitting curve first intersects the spiral F. It follows, therefore, that as the mass of the envelope is increased above the value leading to an infinite radius, no continuous development is possible. What the next stage is in the development of such a star forms a question of great interest. The non-existence of a neighbouring steady state for the models here considered implies

* Provided Z is very small. For $Z = 0{\cdot}10$ the ratio is 1·99.

either that mixing currents will be set up destroying the chemical inhomogeneity or that the star cannot exist in a steady state, i. e. that it will oscillate. In view of the low values of N it is difficult to imagine any mechanism leading to such mixing currents. The second possibility (pulsation) therefore seems to be much more likely, and it is most attractive to identify such pulsating stars with the Cepheid variables.

Whether this peculiar instability to excess mass in the envelope is of practical importance depends on the question of the applicability of the opacity law here assumed (or a sufficiently similar one) and on the occurrence of sufficiently high concentrations of helium in the interior. The minimum helium abundance leading to such an instability depends critically on the opacity law. Roughly speaking, the lower the power of ρ and the higher the power of T in the opacity formula, the lower the minimum value of Y_i. For the opacity law here assumed Y_i has to be at least 0·80 (for $Z=0$) or 0·725 (for $Z=0{\cdot}1$). The mass in the envelope is then 0·386 of the mass of the star for this minimum value of Y_i when the infinite extension is reached, but only 0·07 of the mass for $Y_i=0{\cdot}96$ ($Z=0$) or 0·90 ($Z=0{\cdot}10$).

We consider now values of μ_i/μ_e that are sufficiently small for the fitting curve *not* to intersect the spiral F. In these cases M/M_s decreases monotonically without limit along the fitting curve. The values of r/r_s, however, first drop sharply as the point (S_2, Q_{2e}) moves from the origin along the fitting curve to the region of its closest approach to the spiral F. As soon as the point of closest approach to F has been passed, the value of r/r_s begins to rise again, or at least its drop is sufficiently slow to be counterbalanced by the drop in r/r_s^* along corresponding points of the standard solution. It follows therefore that as the fraction of the total mass in the envelope is increased from zero, the radius of the star increases greatly, reaching a maximum for a fraction corresponding roughly to the closest approach of the fitting curve to F, and thereafter drops again as further mass is added to the envelope. Extremely large radii are therefore associated only with models having a certain moderate fraction of the total mass in the envelope. If more is added (the " redundant masses " as they have been termed by Hoyle and Lyttleton (**2**)), the star shrinks again to a more normal size.

As the mass in the envelope is increased still further, N_{2e} rises more and more and, for sufficiently large mass-ratios $1-M_2/M_s$, this may lead to the onset of convective instabilities either in the envelope or (through its effect on N_{2i}) in the interior. This development is of considerable interest to the theory of stellar evolution, but (since the radii are of more or less normal size) not to the actual red giant problem.

7. So far in our discussion we have compared stars of the same central densities, temperatures and compositions. This is theoretically the simplest approach, but the observationally important characteristics of a star are its mass M_s, radius r_s and its luminosity L. Before conclusions of real significance can be drawn it is therefore necessary to translate to these variables. An exact approach of this kind would be rather complicated, since the law of energy generation would be required to link the luminosity with the central temperature and the central density, and the functions concerned are rather involved. They do, however, have the vital property of making the luminosity depend extremely sensitively on the central temperature and rather insensitively on the density and composition. Accordingly it is permissible, as a first approximation, to treat the central

temperature as a given constant and to carry out higher approximation afterwards, if required. Since T_c is known to vary by less than a factor 5, whereas variations in radius of factors of 100 or more are under investigation, this approximation seems to be permissible. Other writers, in particular Li Hen and Schwarzschild (**3**), have adopted it.

The ratio of the thermal energy per unit mass at the centre of the star to the gravitational potential energy at the surface furnishes a pure number (homology invariant) connecting the mass, radius and central temperature. This invariant will be called D, so that

$$D = \left(\frac{\mathscr{R}T_c}{\mu_i}\right)\left(\frac{r_s}{GM_s}\right). \tag{13}$$

It forms, therefore, a convenient quantity to measure the radial variations of different models. The expression for this invariant in h-variables is

$$\log D = \int_0^{S_a} \frac{Q-S}{S(1+2S-4Q)}\,dS + \log Q_a + \int_{s_a}^{\infty} \frac{N\,dS}{S(1+2S-4Q)}, \tag{14}$$

where the point (S_a, Q_a) is an ***arbitrary*** point on the curve representing the particular model. It is easily proved by differentiation that D is independent of the choice of (S_a, Q_a). For our purposes it will be convenient to choose $(S_a, Q_a) = (S_2, Q_{2i})$. (Note that if (S_a, Q_a) had been chosen to lie outside the discontinuity of ρ, the jump in ρ would have led to an appropriate infinity in N under the last integral in (14).) If expression (14) for D for one of the inhomogeneous models is compared with that for the standard model, then, since $Q_{2i} \doteqdot Q^*(S_2)$ and since $N - N^*$ rapidly approaches zero as S increases from S_2, only the first term in (14) can lead to any appreciable change in D. The ratio D/D^* may therefore be investigated without great difficulty. This important ratio will be called the extension factor and will be denoted by E. It is the ratio of the radius of any model to the radius of a homogeneous star of the same mass, central temperature and central composition. If it is desired to compare the extended model with a main sequence star of the same mass and central temperature, E must be multiplied by μ_i/μ_e.

From what has been said it follows that

$$\log E \doteqdot \int_0^{S_2} \frac{Q-S}{S(1+2S-4Q)}\,dS - \int_0^{S_2} \frac{Q^*-S}{S(1+2S-4Q^*)}\,dS$$
$$= \int_0^{S_2} \frac{(Q-Q^*)(1-2S)\,dS}{S(1+2S-4Q)(1+2S-4Q^*)}. \tag{15}$$

Consideration of the sign of the factors in the last integral shows that as the discontinuity is moved farther into the star along any particular member of the envelope family, E increases if and only if $S<0{\cdot}5$, but decreases for $S>0{\cdot}5$. (It has been assumed that $\mu_i>\mu_e$.) It follows that no larger values of E can occur to the right of $S=0{\cdot}5$ than occur on its left.

In practice, E is most easily calculated by comparing the ratio $Mr_s/M_s r$ for the point (S_2, Q_{2e}) with the corresponding ratio for the standard solution at $S=S_2$. To each point (S_2, Q_{2e}) there corresponds therefore one value of E. Curves of constant E can hence be drawn on the envelope diagram and are shown in Fig. 3. On the spiral F we have $E=\infty$, on the standard solution $E=1$. For any given ratio μ_i/μ_e and any given ratio M_2/M_s the corresponding value of E may be read off the diagram by drawing the fitting curve obtained with fair accuracy by scaling

up the standard curve by the factor μ_i/μ_e. The E value at the point where this curve intersects the curve representing the given ratio M_2/M_s is the extension factor for the model in question.

It will be seen that for any given ratio μ_i/μ_e, as M_2/M_s decreases from unity, either E rises from unity to infinity, and then the model breaks down, or E rises from unity to a maximum and then diminishes again. The first alternative arises if $\mu_i/\mu_e > 2{\cdot}00$, the second if $\mu_i/\mu_e \leqslant 2{\cdot}00$. Curves of E against M_2/M_s are given in Fig. 4 for a number of representative values of μ_i/μ_e. It will be seen that if μ_i/μ_e is less than about 1·8 the maximum extension is only moderate. The representative points of the 15 solutions given in Li Hen and Schwarzschild's paper (**3**) are marked in on Fig. 3. It will be seen that in spite of the extensive calculations, the vital region of high E values has been missed, whereas the approximate analytical methods of Hoyle and Lyttleton (**1**), (**2**) correctly arrive at the existence of such a region.

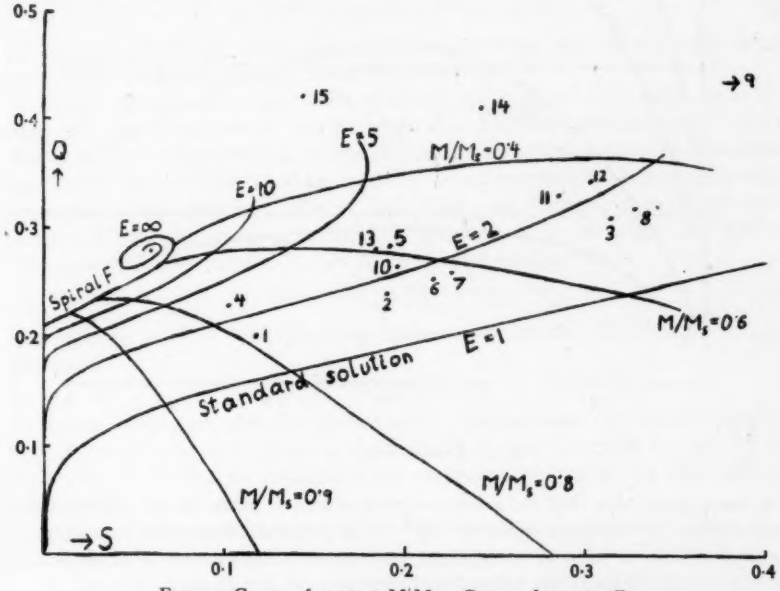


FIG. 3.—*Curves af constant M/M_s. Curves of constant E.*
The numbered dots represent the 15 models of Li Hen and Schwarzschild.

8. The foregoing section deals with the relations between mass and radius, but the luminosity has yet to be considered. For the present opacity law (*cf.* (**4**), Part I, equation (28))

$$\frac{NQ^{8\cdot25}}{S^{1\cdot75}} = \left(\frac{\mathscr{R}}{\mu G}\right)^{7\cdot5}\left(\frac{1}{4\pi}\right)^{2\cdot75}\left(\frac{3\kappa_0 L}{4ac}\right)\left(\frac{r^{1\cdot25}}{M^{5\cdot75}}\right). \tag{16}$$

This equation applies at every point, and in particular when applied at the surface leads to the usual mass-luminosity-radius relation which here takes the form

$$\tfrac{7}{33}A^{8\cdot25} = \left(\frac{\mathscr{R}}{\mu_e G}\right)^{7\cdot5}\left(\frac{1}{4\pi}\right)^{2\cdot75}\left(\frac{3\kappa_{0e}}{4ac}\right)\left(\frac{Lr_s^{1\cdot25}}{M_s^{5\cdot75}}\right) = C \text{ of paper (3).} \tag{17}$$

In the greatly extended models r_s is very much greater than usual, but so is A. In order to obtain a clearer picture of the mass-luminosity relation, it is desirable to eliminate r_s between (13) and (17). This leads to a mass-luminosity-central temperature relation

$$\frac{7}{33}\frac{A^{8\cdot25}}{D^{1\cdot25}}\left(\frac{\mu_e}{\mu_i}\right)^{1\cdot25}=\left(\frac{\mathscr{R}}{\mu_e G}\right)^{6\cdot25}\frac{1}{(4\pi)^{2\cdot75}}\frac{3\kappa_{0e}}{4ac}\frac{L}{T_c^{1\cdot25}M_s^{4\cdot5}}. \tag{18}$$

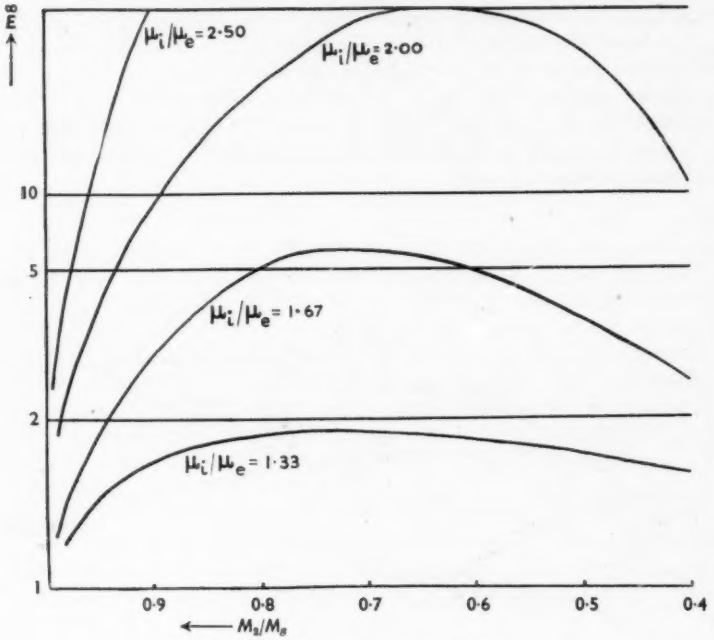


FIG. 4.—*Vertical scale :* $1-E^{-1/2}$.

It is seen from this that in a homologous series of stars of the same central temperature, L is proportional to $M_s^{4\cdot5}$. For practical purposes it is however useful to follow a somewhat different procedure. It can be shown that at an arbitrary point M and r are related to T_c by

$$\frac{\mathscr{R}T_c}{\mu_i}\frac{r}{GM}=\log Q+\int_S^\infty\frac{NdS}{S(1+2S-4Q)}. \tag{19}$$

(Equation (19) may be regarded as a generalization of (14).) Relation (16) is now applied just inside the interface and combined with (19). Then

$$\frac{N_{2i}Q_{2i}^{8\cdot25}}{S_2^{1\cdot75}}\left[\log Q_{2i}+\int_{S_2}^\infty\frac{NdS}{S(1+2S-4Q)}\right]^{-1\cdot25}$$
$$=\left[\left(\frac{\mathscr{R}}{\mu_i G}\right)^{6\cdot25}\frac{1}{(4\pi)^{2\cdot75}}\frac{3\kappa_{0i}}{4ac}\right]\frac{L}{T_c^{1\cdot25}M_2^{4\cdot5}}. \tag{20}$$

It should be observed that the left-hand side of (20) is almost the same whether we consider an extended model or the standard solution. Similarly the

expression in square brackets on the right-hand side is almost the same. Hence $LT_c^{-1\cdot25}M_2^{-4\cdot5}$ is practically the same whether an extended envelope or a homogeneous continuation of the interior is fitted on at the boundary. Assuming the central temperatures to be the same, it follows that

$$L = L^*\left(\frac{M_2}{M_s}\right)^{4\cdot5}\left(\frac{M_2^*}{M_s^*}\right)^{-4\cdot5}, \tag{21}$$

where L^* is the luminosity of a homogeneous star of the same mass, and M_2/M_s, M_2^*/M_s^* are the M/M_s ratios for the extended and homogeneous cases respectively.

If as an example we consider the star of infinite radius that arises for $\mu_i/\mu_e \doteqdot 2\cdot2$ with Q_{2e} on the F spiral at $S_2 = 0\cdot05$, $Q_{2e} = 0\cdot2449$, then since the M/M_s value at this point of the F spiral is 0·679 but equals 0·923 at the corresponding point ($S = 0\cdot05$, $Q = 0\cdot1261$) of the standard solution, the luminosity of such a star will be

$$L = L^*(0\cdot679/0\cdot923)^{4\cdot5} = 0\cdot226L^*, \tag{22}$$

i.e. it is 0·226 of the luminosity of a homogeneous star of the same central temperature and composition and of the same mass.

It may be mentioned that both rule (21) and equation (17) show, as is to be expected, that an extremely thin envelope does not affect the luminosity. In the case of rule (21) this is obvious, but in the case of equation (17) the argument is somewhat intricate and may be regarded as an interesting confirmation of the approximate equation (8). For, a thin layer will leave r_s and M_s effectively unaltered, but will change μ, κ_0 and A. Since close to the origin $Q \doteqdot AS^N$, the change in A will be simply Q_{2e}/Q^*. The change in $\kappa_0\mu^{-7\cdot5}$ is

$$\mu_i^{8\cdot25}N^*/\mu_e^{8\cdot25}N_{2i} = Q_{2e}^{8\cdot25}N^*/Q_{2i}^{8\cdot25}N_{2i}. \tag{23}$$

The change in $A^{8\cdot25}$ will therefore exactly compensate the change in $\kappa_0\mu^{-7\cdot5}$ provided

$$Q^{*8\cdot25}N^* = Q_{2i}^{8\cdot25}N_{2i}. \tag{8'}$$

It can be shown that the terms neglected in (8) are not merely small but actually tend to zero if, with the ratio N/N^* fixed, S is made to tend to zero. Therefore (8′) is exactly satisfied at the surface and hence a very thin layer does not alter the luminosity.

Alternatively this last fact may be considered to be physically obvious and then (8′) can be taken to prove that (8) is exactly satisfied in the limit $S \to 0$.

9. As a final point we consider, supposing the opacity law to remain valid right up to the surface, what would appear to be the actual physical extent of a mathematically infinite (or very large) star.

The observed surface will be not the sphere on which $\rho = 0$ but the sphere from which light reaching us originates on the average. Hence if r_f is the radius of this sphere,

$$\int_{r_f}^{\infty} \kappa\rho dr = 1. \tag{24}$$

Hence

$$\kappa_0\int_{r_f}^{\infty} \rho^{1\cdot75}T^{-3\cdot5}\,dr = 1. \tag{25}$$

In the outer region of the star $N \doteqdot \frac{7}{33}$, so that

$$\rho \sim T^{28/7}.$$

Also for most of the range Q and M are nearly constant, so that $T \sim 1/r$. Hence (25) can be integrated giving

$$\kappa_0 \rho_f^{1\cdot75} T_f^{-3\cdot5} r_f = 2. \tag{26}$$

On the other hand,

$$4\pi a c T_f^4 r_f^2 = L. \tag{27}$$

Finally, it follows from the equation of radiative equilibrium that

$$N = \left(\frac{3\kappa_0 L \mathscr{R}}{16\pi a c \mu G}\right) \frac{\rho_f^{1\cdot75}}{M_f T_f^{6\cdot5}}. \tag{28}$$

Near the surface $N \doteqdot \frac{7}{33}$ and $M_f \doteqdot M_s$.

Combining (26) and (28),

$$\frac{7}{33} = \frac{3L\mathscr{R}}{16\pi a c \mu G M_s} \frac{2}{r_f T_f^3}. \tag{29}$$

Eliminating r between (27) and (29),

$$\frac{7}{33} = \frac{3}{4} \frac{\mathscr{R}}{\mu G M_s} \frac{2}{T_f} \left(\frac{L}{4\pi a c}\right)^{1/2}.$$

This gives for the effective temperature of a very large star

$$T_f = \frac{M_\odot}{M_s} \left(\frac{L}{L_\odot}\right)^{1/2} 10 \text{ deg. K.}$$

This result is so absurdly low compared with observed values that it shows that the actual limitation must almost certainly be due to a breakdown of the opacity formula rather than to the process discussed here. A more significant result is obtained if it is assumed that the opacity drops to zero below a certain temperature, say 2000 deg. Then the effective radius is obtained from the fact that in the outer regions of an infinite star $Q = \frac{7}{33}$. Hence

$$\frac{\mathscr{R} T_f r_f}{G M_s \mu} = \frac{7}{33}.$$

With $T_f = 2000$ deg., $\mu = \frac{1}{2}$ this gives

$$\frac{r_f}{r_\odot} = 1225 \frac{M_s}{M_\odot}.$$

This result assigns a radius of fairly reasonable size to an " infinite " star.

10. In order to interpret the results completely in terms of the theory of stellar evolution it is essential to separate those points of the results which are due to the particular choice of opacity law from those which have some wider significance. It is therefore highly desirable to investigate the same problem, assuming a different opacity law.

Opacity due to electron scattering is independent of ρ and T, and, together with the opacity considered in this paper, is the chief source of opacity in stars. The actual opacity law in any star will be due to some combination of these two effects. We hope to publish shortly an account of an investigation for the case of electron opacity similar to that described in this paper. The results for the two cases may then be expected to bracket the results for the complex opacity law applicable in an actual star, and a more complete view of the problem will then be available.

11. *Conversion of the features of the* 15 *models of Li Hen and Schwarzschild to the terms used in this paper.*

(i) The envelope solutions used by Li Hen and Schwarzschild are described in terms of a constant C connected with the parameter A used in this paper by the relation

$$C=\tfrac{7}{33}A^{8\cdot25}.$$

Models 1, 2, 3, 6, 7, 8, 9 are on the curve $\log_{10}C=-4\cdot6$ $A=0\cdot3343$.
Models 10, 11, 12 are on the curve $\log_{10}C=-4\cdot3$ $A=0\cdot3635$.
Models 4, 5, 13, 14, 15 are on the curve $\log_{10}C=-4\cdot0$ $A=0\cdot3953$.

The first set of models is therefore on a solution intermediate between the standard solution (i) and our solution (iii), the two other sets are between solutions (iii) and (iv). All three sets are very close together.

(ii) The invariants U and V are connected with S and Q by

$$S=U/V,\ Q=1/V.$$

Also
$$N=1/(n+1).$$

(iii) Li Hen and Schwarzschild give their results in the form of the ratio $r_s/r_\odot$ for a star $M_s=10M_\odot$, $T_c=3\times10^7$ deg. K. It is easily calculated from the value of the invariant for the standard solution (Table II (i)) that the radius of a homogeneous star of this mass and central temperature is $8\cdot81\,\mu r_\odot$. Hence the extension factor E is obtained by dividing the ratio $r_s/r_\odot$ given by Li Hen and Schwarzschild by $8\cdot81\,\mu_i$. Table III gives the values of S_2 and Q_{2c} for the 15 models of Li Hen and Schwarzschild, together with the ratio M_2/M_s, the extension factor E and the ratio μ_i/μ_c.

It will be seen from Fig. 3 that the agreement between the figures for the 15 models and our results is most satisfactory. The low E values found by Li Hen and Schwarzschild are to be ascribed purely to the unfortunate choice of A values which lead to a complete absence of points near the spiral F where the models of large radius lie. Model 15 lies well within, and model 14 close to, the border of the part of S-Q plane which cannot be reached by continuously decreasing M_2/M_s (i. e. increasing the mass in the envelope) for fixed μ_i/μ_c, since the fitting curve intersects the spiral F for much higher values of M_2/M_s.

Although Li Hen and Schwarzschild correctly anticipate that higher A values lead to models of much larger radii, they consider such a choice of A to be inappropriate. For, if observational data for red giants are inserted in the formula for A, no very large values result for any plausible composition. To us this argument seems to be invalid. The theory discussed is based on the assumption that the photoelectric effect predominates throughout the star. This is certainly incorrect for very massive stars, as is shown by the fact that the A value computed for massive main sequence stars is extremely small, whereas medium-sized main sequence stars lead to the standard value. But until a complete theory is available, which includes the effects of electron opacity, a simple model must be considered, and theoretical considerations must not be adjusted by premature references to observation. The cardinal observational fact is that the A values of red giants are very much greater than the A values of main sequence stars of similar mass, although both values are depressed by the incidence of electron scattering.

The final suggestion of Li Hen and Schwarzschild that larger radii might be obtained by still further reducing M_2/M_s has been shown in the present paper to be without foundation.

TABLE II

Solutions of the equations for the envelope

(i) *The Standard Solution*

S	Q	N	M/M_s	r/r_s
0	0	0·2121	1	1
0·02	0·1030	0·2139	0·971	0·516
0·03	0·1124	0·2148	0·955	0·481
0·04	0·1198	0·2156	0·939	0·455
0·05	0·1261	0·2165	0·923	0·435
0·06	0·1318	0·2174	0·908	0·418
0·08	0·1419	0·2191	0·878	0·391
0·10	0·1509	0·2210	0·849	0·370
0·15	0·1712	0·2257	0·781	0·332
0·20	0·1896	0·2303	0·721	0·305
0·25	0·2071	0·2347	0·668	0·285
0·30	0·2241	0·2391	0·621	0·269
0·40	0·2584	0·2478	0·541	0·243
0·50	0·2918	0·2562	0·478	0·225
0·60	0·3245	0·2645	0·426	0·210
0·70	0·3568	0·2727	0·382	0·199
0·80	0·3891	0·2807	0·346	0·190
1·00	0·4538	0·2963	0·289	0·174
1·2	0·5188	0·3115	0·247	0·1625
1·4	0·5841	0·3263	0·214	0·1525
1·6	0·6497	0·3406	0·187	0·1447
1·8	0·7156	0·3548	0·166	0·1379
2·0	0·7817	0·3689	0·149	0·1320
2·2	0·8481	0·3827	0·134	0·1266
2·4	0·9147	0·3963	0·122	0·1220
2·45	0·9314	0·4000	0·119	0·1210

$A = 0{\cdot}256$

Radius of core in Emden units
$x = 1{\cdot}120$

$$\frac{r_s}{GM_s}\frac{\mathscr{R}T_c}{\mu_c} = 1{\cdot}15$$

$$\frac{\rho_{\text{mean}}}{\rho_c} = 0{\cdot}0123$$

Radius of star in Emden units
$x_s = 9{\cdot}25$

(ii) *The Spiral F*

S	Q	N	M/M_s
0	0·2121	0·2121	1·000
0·01	0·2188	0·2132	0·935
0·02	0·2258	0·2143	0·871
0·03	0·2331	0·2154	0·808
0·04	0·2409	0·2166	0·744
0·05	0·2494	0·2177	0·679
0·06	0·2589	0·2188	0·610
0·07	0·2702	0·2199	0·528

$r/r_s = 0$ along the spiral

$A = \infty$

$E = \infty$

The spiral continues, approaching (after an infinity of convolutions) the point $S = \frac{5}{82} \doteqdot 0{\cdot}0610$, $Q = \frac{23}{82} \doteqdot 0{\cdot}2805$, $N = \frac{9}{41} \doteqdot 0{\cdot}2195$.

TABLE II—*continued*

(iii) *The Solution A*=0·354

S	*Q*	*N*	M/M_s	r/r_s	*E*
0	0	0·2121	1	1	1
0·02	0·1357	0·2140	0·966	0·364	1·41
0·04	0·1581	0·2160	0·926	0·290	1·54
0·06	0·1734	0·2179	0·885	0·248	1·65
0·08	0·1875	0·2199	0·843	0·219	1·72
0·10	0·2000	0·2218	0·803	0·197	1·78
0·15	0·2301	0·2270	0·706	0·156	1·93
0·20	0·2605	0·2323	0·617	0·129	2·03
0·25	0·2993	0·2377	0·533	0·107	2·12
0·30	0·3306	0·2435	0·451	0·090	2·16
0·3215	0·350	0·2462	0·415	0·082	2·20
0·3443	0·375	0·2493	0·373	0·0736	2·24
0·3596	0·40	0·2520	0·335	0·0646	2·29
0·3641	0·41	0·2529	0·321	0·0626	2·30
0·3672	0·42	0·2537	0·308	0·0586	2·32
0·3688	0·43	0·2544	0·294	0·0558	2·34
0·3687	0·44	0·2548	0·281	0·0530	2·36
0·3667	0·45	0·2553	0·269	0·0500	2·38
0·3623	0·46	0·2556	0·256	0·0470	2·41
0·3550	0·47	0·2556	0·243	0·0440	2·44

(iv) *The Solution A*=0·442

S	*Q*	*N*	M/M_s	r/r_s	*E*
0	0	0·2121	1	1	1
0·0050	0·1262	0·2126	0·993	0·411	1·57
0·0075	0·1359	0·2129	0·989	0·369	1·66
0·010	0·1432	0·2131	0·983	0·336	1·73
0·015	0·1542	0·2136	0·972	0·292	1·85
0·02	0·1625	0·2141	0·959	0·261	1·96
0·03	0·1755	0·2151	0·934	0·217	2·17
0·04	0·1860	0·2161	0·907	0·186	2·36
0·06	0·2041	0·2183	0·853	0·145	2·71
0·08	0·2202	0·2204	0·797	0·118	3·00
0·10	0·2360	0·2227	0·740	0·098	3·30
0·12	0·2524	0·2250	0·678	0·081	3·60
0·14	0·2700	0·2275	0·618	0·068	3·89
0·15	0·2800	0·2287	0·586	0·0615	4·05
0·1676	0·3000	0·2308	0·524	0·0502	4·4
0·1800	0·3200	0·2327	0·467	0·0407	4·9
0·1860	0·3400	0·2341	0·413	0·0323	5·4
0·1800	0·3600	0·2343	0·354	0·0238	6·3
0·1750	0·3650	0·2340	0·336	0·0212	6·7

TABLE II—*continued*

(v) *The Solution A*=0·590

S	Q	N	M/M_s	r/r_s	E
0	0	0·2121	1	1	1
0·0025	0·1400	0·2123	0·995	0·343	2·0
0·005	0·1570	0·2126	0·989	0·266	2·4
0·01	0·1750	0·2131	0·975	0·191	3·0
0·02	0·1950	0·2142	0·941	0·120	4·2
0·025	0·2018	0·2148	0·923	0·104	4·8
0·03	0·2078	0·2153	0·905	0·088	5·3
0·04	0·2186	0·2165	0·861	0·066	6·4
0·05	0·2285	0·2176	0·820	0·051	7·6
0·06	0·2381	0·2187	0·775	0·0403	8·8
0·07	0·2481	0·2199	0·727	0·0316	10·0
0·0718	0·2500	0·2201	0·717	0·0301	10·5
0·0812	0·26	0·2213	0·668	0·0238	11·5
0·0898	0·27	0·2224	0·617	0·0188	12·0
0·0968	0·28	0·2233	0·569	0·0147	13·0
0·1021	0·29	0·2241	0·518	0·0113	15·0
0·1048	0·30	0·2246	0·465	0·0083	20·0
0·1047	0·305	0·2247	0·436	0·0069	24·0
0·1029	0·310	0·2247	0·402	0·0054	32·0
0·0975	0·3125	0·2246	0·378	0·0045	37·0

TABLE III

Model	μ_i/μ_e	M_2/M_s	S_2	Q_{2c}	E
1	1·33	0·780	0·118	0·200	1·475
2	1·35	0·659	0·191	0·238	1·560
3	1·40	0·489	0·315	0·309	1·685
4	1·51	0·734	0·118	0·229	2·285
5	1·58	0·572	0·191	0·282	2·645
6	1·29	0·620	0·217	0·253	1·610
7	1·28	0·600	0·227	0·260	1·630
8	1·36	0·472	0·328	0·318	1·680
9	1·57	0·305	0·463	0·432	1·745
10	1·40	0·610	0·197	0·263	1·990
11	1·51	0·453	0·288	0·330	2·175
12	1·54	0·423	0·305	0·346	2·205
13	1·52	0·590	0·182	0·275	2·580
14	2·06	0·281	0·244	0·414	3·745
15	2·55	0·200	0·144	0·421	5·42

Newnham College, Cambridge:
1950 *May* 3.

Trinity College, Cambridge.

References

(1) F. Hoyle and R. A. Lyttleton, *M.N.*, **102**, 218, 1942.
(2) F. Hoyle and R. A. Lyttleton, *M.N.*, **109**, 605, 1950.
(3) Li-Hen and M. Schwarzschild, *M.N.*, **109**, 631, 1950.
(4) C. M. Bondi and H. Bondi, *M.N.*, **109**, 62, 1949.
(5) C. M. Bondi, *M.N.*, **110**, 275, 1950.

SELF-GRAVITATING STAR SYSTEMS

G. L. Camm

(Received 1949 November 23)

Summary

The Boltzmann equation for a star system in which the controlling forces are solely due to the gravitational attractions between the members is not strictly linear. For the restricted case of a system of stars of equal mass, stratified in plane parallel layers, the Boltzmann equation reduces to a non-linear partial differential equation in three independent variables. Three functions which satisfy this equation have been found. Two of these correspond to the equilibrium configurations of an isothermal self-gravitating gas, and of an adiabatic gas. The hypothetical models can be compared with the greatly flattened galactic system, and there is considerable agreement between the observed variation of density with distance from the galactic plane and the variation predicted by one of the three solutions.

---

1. *General theory.*—The distribution in space and the distribution of velocities for a star system at time t can be described mathematically by the velocity-distribution function. If we adopt any convenient non-accelerated Cartesian frame of space-coordinates, x, y, z, and corresponding velocity-components, u, v, w, then the velocity-distribution function, $f(t, x, y, z, u, v, w)$, is such that the number of stars at time t in a volume $dx\,dy\,dz$ surrounding the point (x, y, z) for which the velocity-components lie in the ranges u to $u+du$, v to $v+dv$, w to $w+dw$ is the quantity, $f(t, x, y, z, u, v, w)\ dxdydzdudvdw$. The function f is not completely arbitrary, since it must satisfy a generalized Boltzmann equation of continuity, namely,

$$\frac{\partial f}{\partial t} + u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} + \frac{du}{dt}\frac{\partial f}{\partial u} + \frac{dv}{dt}\frac{\partial f}{\partial v} + \frac{dw}{dt}\frac{\partial f}{\partial w} = 0.$$

The acceleration, of which the three components are $du/dt, dv/dt, dw/dt$, is produced by the gravitational attractions of the rest of the stars in the system. The frequency of close encounters between two stars whereby their orbits are appreciably altered depends on the local values of the star-density and the velocity-dispersion (**1**). In the region of the Sun, and probably throughout most of the galactic system, the frequency of such encounters is extremely small. It is usual, therefore, in considering the dynamics of stellar systems to ignore completely the possible effects of individual encounters. However, it is clear that the *stability* of a star system is likely to be affected by encounters, though this problem is outside the scope of the present paper.

Ignoring encounters, we assume that the gravitational forces can be derived from a potential function, $\Omega(t, x, y, z)$, which is the gravitational potential of the smoothed-out distribution of mass. The equation of continuity then takes the form

$$\frac{\partial f}{\partial t} + u\frac{\partial f}{\partial x} + v\frac{\partial f}{\partial y} + w\frac{\partial f}{\partial z} + \frac{\partial \Omega}{\partial x}\frac{\partial f}{\partial u} + \frac{\partial \Omega}{\partial y}\frac{\partial f}{\partial v} + \frac{\partial \Omega}{\partial z}\frac{\partial f}{\partial w} = 0. \qquad (1)$$

This equation has been treated in two distinct ways; one method, due to Eddington (**2**), is to assume the form in which f contains the variables u, v, w, and hence, from

(1), to find f as a function of t, x, y, z, u, v, w, and Ω as a function of t, x, y, z. This method has not so far yielded any solution which corresponds satisfactorily to the observed distribution of velocity and density (**3**), (**4**), (**5**). The alternative method, due to Jeans (**6**), is to recognize (1) as a linear partial differential equation in f, with corresponding auxiliary equations, of which there are six integrals. Only two such integrals have been found, and these only in the special circumstances of a steady state with rotational symmetry. In this case the total energy and the angular momentum are integrals. Unfortunately these two integrals alone can never give rise to a velocity-ellipsoid with three unequal axes, which is the observed feature of the velocity-distribution of stars in the galactic system.

However, the chief objection to this second approach is that (1) is only apparently a linear differential equation in f. This is most obvious if we consider two systems of stars in each of which all the stars have the same mass. Suppose that the velocity-distribution functions are f_1 and f_2, and that at $t=0$, $f_1(t, x, y, z, u, v, w) = 2f_2(t, x, y, z, u, v, w)$, for all x, y, z, u, v, w. If (1) were truly linear, then we should have $f_1 = 2f_2$ for all t. Actually, since there are twice as many stars in the first system, we must have at $t=0$, $\Omega_1(t, x, y, z) = 2\Omega_2(t, x, y, z)$. Consequently the Boltzmann equation satisfied by f_1 is not simply a multiple of that satisfied by f_2.

In the galactic system it seems that stars of different spectral types have different distributions of velocity. Each type has its own function f, satisfying (1), but the components of gravitational attraction occurring in (1) are independent of spectral type. Nevertheless (1) is still not a linear equation in f, for, as the function f for any one type varies, so the total star-density will change, and the gravitational potential will also be affected. It therefore seems desirable to attempt a non-linear treatment of the problem for stars of equal mass—though it is recognized that the mathematical difficulties of non-linear differential equations are such that our primary aim must be to find any solution at all. Even for the linear problem the number of solutions so far discovered is extremely small.

The problem is obviously similar to that of a self-gravitating gas in hydrostatic equilibrium. In both, the effects of collisions are ignored. In star systems the density is so low that collisions are very infrequent; the system is still in such an early stage of development that collisions have so far played a negligible part. On the other hand, the relaxation time for a gas, unless under extraordinary conditions of temperature and density, is so short that it rapidly reaches a steady state in which collisions occur, but do not affect the statistical distribution. The comparison between the stellar system and a self-gravitating gas extends to the equation relating " pressure " with the field of gravitational force. Equation (1) is obtained from considerations of continuity, but it is quite analogous to the hydrostatic equation connecting p, the pressure, and ρ, the density, with Ω, namely,

$$\nabla p = \rho \nabla \Omega.$$

In the case of the gas there is also a relation, $p \propto \rho^{\gamma}$, connecting pressure and density. This relation depends on the change in internal energy of the molecules due to collisions. Clearly there is no analogous relation in a star system. However, it will be shown that certain solutions of our problem, solutions selected for their mathematical simplicity, yield distributions of density which correspond exactly to the self-gravitating gas.

2. *The case of a stratified star system.*—The simplest case we can consider is that of a star-system which is stratified in plane parallel layers. The velocity distribution is independent of the space variables x and y, and it follows that the density and the gravitational potential are also independent of x and y. We shall also assume in the first instance that the system is in a steady state, so that t is not explicitly involved. This is purely for convenience in describing the method, which can be applied equally well to non-steady systems. However, since no non-steady systems of physical interest have been found, the explicit retention of the variable t seems unnecessary. The function $f(z, w)$ is now defined so that $f(z, w)\,dz\,dw$ is the number of stars in a column of unit cross-section extending from a height z to $z+dz$ for which the components of velocity in the z-direction lie between w and $w+dw$. Then f has to satisfy the equation

$$w\frac{\partial f}{\partial z}+\frac{d\Omega}{dz}\frac{\partial f}{\partial w}=0. \tag{2}$$

We also have that $\nu(z)$, the star-density, is given by the expression

$$\nu(z)=\int_{-\infty}^{+\infty} f(z, w)\,dw.$$

From this we obtain $\rho(z)$, the mass-density

$$\rho(z)=m\int_{-\infty}^{+\infty} f(z, w)\,dw. \tag{3}$$

Finally, Poisson's equation takes the form

$$\frac{d^2\Omega}{dz^2}=-4\pi G\rho(z). \tag{4}$$

These three equations, (2), (3) and (4), can be combined into a single equation if we introduce two new functions in the following way. Let us define $F(z, s)$ by the following equation:

$$F(z, s)=m\int_{-\infty}^{+\infty} f(z, w)e^{iws}\,dw;$$

further, let us put

$$K(z, s)=2\pi G\left[\int_{z}^{+\infty} F(\lambda, s)\,d\lambda-\int_{-\infty}^{z} F(\lambda, s)\,d\lambda\right].$$

It is seen that $F(z, 0)$ is the mass-density, and $K(z, 0)$ is the gravitational attraction at the level z. Now if we multiply (2) by me^{iws} and integrate with respect to w from $-\infty$ to $+\infty$, we have the equation

$$m\frac{\partial}{\partial z}\left\{\int_{-\infty}^{+\infty} wf(z, w)e^{iws}\,dw\right\}+K(z, 0)m\int_{-\infty}^{+\infty}\frac{\partial f}{\partial w}e^{iws}\,dw=0.$$

In terms of F this can be written in the form

$$-i\frac{\partial^2}{\partial s\partial z}F(z,s)-is\,K(z,0)\,F(z,s)=0,$$

and in terms of $K(z,s)$ it becomes

$$\frac{\partial^3 K(z,s)}{\partial s\partial z^2}+sK(z,0)\frac{\partial K(z,s)}{\partial z}=0. \tag{5}$$

This is the general equation which must be satisfied for every possible equilibrium configuration of a stratified self-gravitating star system. There is no other

restriction on $K(z, s)$ except that it must be such a function as to make $f(z, w)$ non-negative.

For the system in a non-steady state, the distribution is completely specified if we determine the function $K(t, z, s)$ which satisfies the equation

$$\frac{\partial^2 K(t,z,s)}{\partial t \partial z} - i\frac{\partial^3 K(t,z,s)}{\partial s \partial z^2} - is\,K(t,\,z,\,0)\frac{\partial K(t,z,s)}{\partial z} = 0.$$

3. *A particular solution for the stratified system.*—To obtain one solution of (5) we assume that, although the star-density varies with z, yet at any level the percentage of stars with any selected velocity is quite independent of the level. In terms of the velocity-distribution function, we assume that f is simply the product of a function of z and a function of w. In (5) we write

$$K(z, s) = Z(z)\,S(s),\ \ S(0) = 1.$$

Thus we find that Z and S satisfy the equation

$$\frac{dS}{ds}\frac{d^2Z}{dz^2} + sZ\frac{dZ}{dz}S = 0.$$

Separating the variables, we obtain

$$\frac{dS}{ds} = -h^2 sS, \quad h^2\frac{d^2Z}{dz^2} = Z\frac{dZ}{dz}, \tag{6),(7}$$

where h is an arbitrary constant. From (6) we obtain $S = \exp(-\frac{1}{2}h^2s^2)$, since we also require that $S(0) = 1$. This gives

$$F(z,\ s) = -\frac{1}{4\pi G}\frac{dZ}{dz}\exp\left(-\tfrac{1}{2}h^2s^2\right).$$

It therefore follows, by Fourier's integral theorem, that

$$f(z,\ w) = \frac{1}{2\pi m}\int_{-\infty}^{+\infty} F(z,\ s)e^{-iws}\,ds$$

$$= -\frac{1}{8\pi^2 Gm}\frac{dZ}{dz}\int_{-\infty}^{+\infty}\exp\left(-\tfrac{1}{2}h^2s^2 - iws\right)ds.$$

The integral converges only if $h^2 > 0$, so that the sign of that arbitrary constant is the correct one. The function f is therefore given by

$$f(z, w) = -\frac{1}{8\pi^2 Gm}\frac{dZ}{dz}\int_{-\infty}^{+\infty}\exp\{-\tfrac{1}{2}(hs + iw/h)^2 - \tfrac{1}{2}w^2/h^2\}\,ds$$

$$= -\frac{2^{1/2}}{8\pi^{3/2}Gmh}\frac{dZ}{dz}\exp(-\tfrac{1}{2}w^2/h^2). \tag{8}$$

The function Z is now determined from (7), which integrates to give

$$2h^2\frac{dZ}{dz} = Z^2 + \text{constant}. \tag{9}$$

Now since $dZ/dz = -4\pi G\rho$, dZ/dz is essentially negative, so that the constant is negative, $-C^2$ say, and Z always lies between $-C$ and $+C$. Let us put $Z = -C\tanh\theta$, so that $dZ/dz = -C\,\mathrm{sech}^2\theta\,d\theta/dz$. On substituting in (9) we obtain the equation

$$-2h^2C\,\mathrm{sech}^2\theta\frac{d\theta}{dz} = -C^2 + C^2\tanh^2\theta,$$

which can be reduced to give

$$\frac{d\theta}{dz} = C/2h^2.$$

If z is measured from the level at which Z vanishes, then we must have $\theta = Cz/2h^2$. It now follows that $Z = -C\tanh(Cz/2h^2)$, and $\rho = (C^2/8\pi Gh^2)\operatorname{sech}^2(Cz/2h^2)$. Thus the density reaches a maximum value, $\rho_0 = C^2/8\pi Gh^2$, in the plane $z = 0$. It is convenient to specify the distribution in terms of the two parameters, ρ_0, the maximum density, and h, the root-mean-square velocity at the level of maximum density. In terms of these quantities we have

$$\begin{aligned}\frac{d\Omega}{dz} &= Z(z)\\ &= -2h(2\pi G\rho_0)^{1/2}\tanh\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\},\\ \rho(z) &= \rho_0\operatorname{sech}^2\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\},\\ f(z,w) &= \frac{\rho_0}{(2\pi)^{1/2}mh}\operatorname{sech}^2\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}\exp(-\tfrac{1}{2}w^2/h^2).\end{aligned}$$

The total mass in a column of unit cross-section, denoted by M, is given by the expression

$$\begin{aligned}M &= \int_{-\infty}^{+\infty}\rho_0\operatorname{sech}^2\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}dz\\ &= (2h^2\rho_0/\pi G)^{1/2}.\end{aligned}$$

The gravitational potential at any level can also be determined. We have the relation

$$\frac{d\Omega}{dz} = -2h(2\pi G\rho_0)^{1/2}\tanh\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}.$$

This equation will integrate, to give

$$\Omega = -2h^2\log\cosh\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\},$$

where we have arranged to make Ω vanish in the plane $z = 0$. It follows that

$$e^{\Omega/h^2} = \operatorname{sech}^2\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}.$$

This can be substituted in the expression for $f(z,w)$, which can then be written

$$f(z,w) = \frac{\rho_0}{(2\pi)^{1/2}mh}\exp\{-(\tfrac{1}{2}w^2 - \Omega)/h^2\}.$$

The solution we have found is therefore one in which the velocity-distribution is a function of the total energy only. This is not entirely surprising, for we have assumed that there is a steady state in which there is a steady gravitational field, $d\Omega/dz$; thus equation (2) is effectively a linear equation in f, and, as such, its solution is bound to be of the form, $f\{\frac{1}{2}w^2 - \Omega(z)\}$.

A solution of this kind, in which the Gaussian distribution of velocities was assumed, was given by Oort (**7**). He did not use the property that the system was a self-gravitating one, but determined the gravitational force as a function of height by estimating the variation of density with distance from the galactic plane. A more detailed comparison with Oort's work, and especially with his

analysis of the observational data, will be given later in this paper. An approximate solution, valid near the galactic plane, has been given by Chandrasekhar (**8**).

One further remark should be made about this distribution of density. It is precisely the equilibrium configuration of an isothermal gas, stratified in parallel layers, and subject only to its own gravitational attraction. Our mathematical assumption that the proportion of stars with any selected velocity was independent of height corresponds exactly to the assumption in the gaseous problem that the gas is isothermal.

Although attention has been confined to the determination of a steady configuration, it is possible to investigate by the same method the existence of non-steady states. If we assume that the velocity-distribution function $f(t, z, w)$ can be written as a product, $Z(z, t)\ W(w)$, then the only result we can obtain is the expression

$$f(t,z,w)=\frac{\rho_0}{(2\pi)^{1/2}mh}\operatorname{sech}^2\left\{\frac{z-ut}{h}(2\pi G\rho_0)^{1/2}\right\}\exp\{-\tfrac{1}{2}(w-u)^2/h^2\}.$$

This is not a new solution, but is essentially the distribution previously obtained for the steady state, but now moving uniformly with a constant speed u in the z-direction.

4. *A more general solution.*—Although the complete solution of the problem has not been found, a more general solution than that just given can be obtained. Again the working will be given only for the steady problem. It can in fact be shown that there are no solutions of the type now to be considered in which the time is explicitly involved. In the previous section it was assumed that the proportion of stars with any given velocity was independent of position; this condition is now replaced by a similar but less stringent one. We suppose that in the function $f(z, w)$, w always occurs with a factor, $s(z)$, a function of z still to be determined. We shall therefore write $f(z, q)$, where q is the second independent variable, being the product, $ws(z)$. Finally, we assume that $f(z, q)$ is the product of two functions, $Z(z)$ and $Q(q)$. It should be noticed that $s(z)$ may not vanish except at a level where f itself vanishes—otherwise infinite values of w would be possible. We therefore take $s(z)$ to be positive, or possibly zero at isolated levels. It must also be emphasized that in the first instance we will confine attention to systems in which very large values of w can occur, even though they may be highly improbable. This implies that there is no limit to the total energy a particle may possess, and consequently the spatial distribution of matter will extend to infinity.

Since only a special kind of solution is being considered, there is nothing to be gained by the use of Fourier transforms, and we proceed directly. From (3) we have the equation

$$\begin{aligned}\rho(z)&=mZ(z)\int_{w=-\infty}^{+\infty}Q(ws)\,dw\\&=m\frac{Z(z)}{s(z)}\int_{q=-\infty}^{+\infty}Q\,dq.\end{aligned}$$

The integral is a constant, and we choose Q so as to make the constant unity. Now let us put $d\Omega/dz=K(z)$, so that

$$\frac{dK}{dz}=-4\pi GmZ/s.$$

It follows that

$$Z = -\frac{1}{4\pi Gm} s \frac{dK}{dz},$$

and also that

$$f(z, w) = -\frac{1}{4\pi Gm} s \frac{dK}{dz} Q(sw).$$

Hence we have

$$\frac{\partial f}{\partial w} = -\frac{1}{4\pi Gm} s^2 \frac{dK}{dz} \frac{dQ}{dq}$$

and

$$\frac{\partial f}{\partial z} = -\frac{1}{4\pi Gm} \left\{ \frac{d}{dz}\left(s\frac{dK}{dz}\right) Q + \left(s\frac{dK}{dz}\right) \frac{q}{s} \frac{ds}{dz} \frac{dQ}{dq} \right\}.$$

These expressions are substituted in (2), and after some cancellation we obtain the equation

$$\frac{q}{s} \frac{d}{dz}\left(s\frac{dK}{dz}\right) Q + \frac{q^2}{s} \frac{dK}{dz} \frac{ds}{dz} \frac{dQ}{dq} + s^2 K \frac{dK}{dz} \frac{dQ}{dq} = 0. \tag{10}$$

Dividing through by dQ/dq, and differentiating twice with respect to q, we obtain

$$\frac{1}{s} \frac{d}{dz}\left(s\frac{dK}{dz}\right) \frac{d^2}{dq^2}\left\{ qQ \Big/ \frac{dQ}{dq} \right\} + \frac{2}{s} \frac{dK}{dz} \frac{ds}{dz} = 0. \tag{11}$$

Clearly $d/dz\{s(dK/dz)\}$ cannot be identically zero, since it would imply that the density was constant; it follows that

$$\frac{d^2}{dq^2}\left\{ qQ \Big/ \frac{dQ}{dq} \right\} = \text{constant} = 2l.$$

On integration, this gives

$$qQ \Big/ \frac{dQ}{dq} = lq^2 + 2mq + n, \tag{12}$$

where m and n are arbitrary constants. From (11) it is seen that

$$l \frac{d}{dz}\left(s\frac{dK}{dz}\right) = -\frac{dK}{dz} \frac{ds}{dz}. \tag{13}$$

From (10) it follows that

$$\frac{d}{dz}\left(s\frac{dK}{dz}\right)(lq^2 + 2mq + n) + q^2 \frac{dK}{dz} \frac{ds}{dz} + s^3 K \frac{dK}{dz} = 0.$$

As this is an identity in q, we must have that m is identically zero, and also that

$$n \frac{d}{dz}\left(s\frac{dK}{dz}\right) + s^3 K \frac{dK}{dz} = 0. \tag{14}$$

Equation (12) can now be written in the form

$$\frac{1}{Q} \frac{dQ}{dq} = \frac{q}{lq^2 + n}.$$

If $l = 0$, then Q varies as $e^{q^2/2n}$. This case needs no further consideration, for it implies $ds/dz = 0$, and therefore reproduces the solution we have previously obtained, in which the velocity-dispersion is independent of position. If $l \neq 0$, then $Q = A(q^2 + n/l)^{1/2l}$. Since we are not imposing any limit on w, q can take all values; clearly Q must tend to zero as q approaches infinity, so that the constant l must be negative. We must also distinguish between the cases in which n/l is

positive, negative, or zero. If n/l is negative, $-a^2$, say, then $Q \propto (q^2-a^2)^{1/2l}$, and we find that Q is infinite for $q = \pm a$. If n is zero, then Q is infinite for $q = 0$. On physical grounds Q must never be infinite. If $n/l = a^2$, then we have

$$Q = A(q^2+a^2)^{1/2l}.$$

It is seen that Q is finite for all q, and tends to zero as q tends to $\pm\infty$. The constant A is determined by the condition

$$1 = A\int_{-\infty}^{+\infty}(q^2+a^2)^{1/2l}\,dq.$$

This integral will converge only if $1/l < -1$. The mean value of q exists (and is zero) if $1/l < -2$, and the mean value of q^2 exists if $1/l < -3$. Its value is readily found, for we have

$$\overline{q^2} = A\int_{-\infty}^{+\infty} q^2(q^2+a^2)^{1/2l}\,dq$$

$$= \left[\frac{Aq}{2+1/l}(q^2+a^2)^{1+1/2l}\right]_{-\infty}^{+\infty} - \frac{A}{2+1/l}\int_{-\infty}^{+\infty}(q^2+a^2)^{1+1/2l}dq.$$

It therefore follows that

$$(2+1/l)\overline{q^2} = -A\int_{-\infty}^{+\infty}(q^2+a^2)^{1+1/2l}dq$$

$$= -A\int_{-\infty}^{+\infty} q^2(q^2+a^2)^{1/2l}dq - Aa^2\int_{-\infty}^{+\infty}(q^2+a^2)^{1/2l}dq$$

$$= -\overline{q^2} - a^2,$$

or
$$\overline{q^2} = -a^2/(3+1/l).$$

For convenience, we can write

$$3+1/l = -2n, \quad n>0.$$

From (13) and (14) we obtain

$$a^2\frac{dK}{dz}\frac{ds}{dz} = s^3K\frac{dK}{dz},$$

or

$$a^2\frac{ds}{dz} = s^3K. \tag{15}$$

Equation (13) can be re-written as

$$l\frac{d}{dz}\left\{\log\left(s\frac{dK}{dz}\right)\right\} = -\frac{d}{dz}(\log s),$$

and it can be integrated, to give

$$(l+1)\log s + l\log\frac{dK}{dz} = \text{constant},$$

or
$$\frac{dK}{dz} = -Cs^{-(1+1/l)} = -Cs^{2(n+1)}. \tag{16}$$

Here, C is a positive arbitrary constant; the negative sign is included since dK/dz is essentially negative. From (15) and (16) it is seen that

$$s^3K\frac{dK}{dz} = -Ca^2s^{2(n+1)}\frac{ds}{dz},$$

or
$$2K\frac{dK}{dz} = -2Ca^2s^{2n-1}\frac{ds}{dz}.$$

This can be integrated, and gives the equation

$$K^2 = -\frac{C}{n}a^2 s^{2n} + \text{constant}.$$

Now K vanishes at the level which divides a column of unit cross-section into two equal masses. This is taken to be the level $z=0$; at that level $s(z)$ takes the value b. Hence we have the equation

$$K^2 = \frac{C}{n}a^2(b^{2n} - s^{2n}).$$

Since C and n are positive, b must be the maximum value of $s(z)$, attained in the plane $z=0$, where K vanishes. It is therefore convenient to write

$$s^{2n} = b^{2n}\cos^2\theta,$$

or

$$s = b\cos^{n^{-1}}\theta,$$

where θ is a function of z, vanishing at $z=0$.

This gives

$$K^2 = \frac{C}{n}a^2 b^{2n}\sin^2\theta,$$

or

$$K = -\left(\frac{C}{n}\right)^{1/2} ab^n \sin\theta;$$

the negative sign is taken, since K is negative when z is positive. Differentiating with respect to z, we obtain

$$\frac{dK}{dz} = -\left(\frac{C}{n}\right)^{1/2} ab^n \cos\theta\frac{d\theta}{dz},$$

and from (16) we have

$$Cb^{2n+2}\cos^{2(1+n^{-1})}\theta = \left(\frac{C}{n}\right)^{1/2} ab^n\cos\theta\frac{d\theta}{dz}.$$

This gives z as an integral,

$$z = \frac{1}{(Cn)^{1/2}}ab^{-n-2}\int_0^\theta \sec^{1+2n^{-1}}\phi\, d\phi. \tag{17}$$

Since the index of $\sec\phi$ is greater than unity, z tends to infinity as θ tends to $\pi/2$, and the density tends to zero.

The density has a maximum value, ρ_0, at the level $z=0$, where we have

$$\rho_0 = \frac{C}{4\pi G}b^{2n+2}.$$

The general value of the density is given by

$$\rho = \rho_0\cos^{2(1+n^{-1})}\theta.$$

The total mass in a column of unit cross-section is M, where we have

$$M = \operatorname*{Lt}_{z\to\infty}\{-K(z)/2\pi G\} = \left(\frac{C}{n}\right)^{1/2}\frac{ab^n}{2\pi G}.$$

This could be calculated directly, as it is also given by the integral

$$\begin{aligned} M &= 2\int_{z=0}^{\infty}\rho\, dz \\ &= 2\rho_0\int_{\theta=0}^{\pi/2}\cos^{2(1+n^{-1})}\theta\frac{dz}{d\theta}d\theta \\ &= \frac{2\rho_0}{(Cn)^{1/2}}ab^{-n-2}\int_0^{\pi/2}\cos\theta d\theta. \end{aligned}$$

But we also have

$$C = 4\pi G\rho_0 b^{-2n-2},$$

so that we can write

$$M = (a^2\rho_0/\pi G n b^2)^{1/2}.$$

The mean square velocity, $\overline{w^2}$ varies with z, its value being $\overline{q^2}/s^2$. Now we have seen that $\overline{q^2} = a^2/2n$, so that h^2, the mean square velocity in the plane $z = 0$, is $a^2/2nb^2$. This is substituted in the expression for M, which can then be written

$$M = (2\rho_0 h^2/\pi G)^{1/2}.$$

This formula was also obtained in the case of the "isothermal" distribution, considered in the previous section.

The expression for the velocity-distribution is now given by the equation

$$f(z, w) = \frac{\rho_0}{m} s(z) \cos^{2(1+n^{-1})}\theta \; A(q^2 + a^2)^{-n-3/2},$$

or, if we express $s(z)$ in terms of θ,

$$f(z, w) = \frac{\rho_0 b}{m} \cos^{2+3n^{-1}}\theta \; A(q^2 + a^2)^{-n-3/2}. \tag{18}$$

The value of A is found from the condition

$$1 = A \int_{-\infty}^{+\infty} (q^2 + a^2)^{-n-3/2}\, dq.$$

Putting $q = ax$, we have

$$A^{-1} = 2a^{-2(n+1)} \int_0^\infty (1 + x^2)^{-n-3/2}\, dx = a^{-2(n+1)}\pi^{1/2} n!/(n + \tfrac{1}{2})!.$$

This result is substituted in (18) which then becomes

$$f(z, w) = \frac{\rho_0 b}{m\pi^{1/2}} \frac{(n + \frac{1}{2})!}{n!} a^{2(n+1)} \cos^{2+3n^{-1}}\theta (q^2 + a^2)^{-n-3/2}.$$

We also have that $q = ws$, or $wb\cos^{n^{-1}}\theta$, so that f can also be expressed by the equation

$$\begin{aligned} f(z, w) &= \frac{\rho_0 b}{m\pi^{1/2}} \frac{(n + \frac{1}{2})!}{n!} a^{2n+2} \cos^{2+3n^{-1}}\theta \; (a^2 + w^2 b^2 \cos^{2n^{-1}}\theta)^{-n-3/2} \\ &= \frac{\rho_0}{m\pi^{1/2}} \frac{(n + \frac{1}{2})!}{n!} \left(\frac{a}{b}\right)^{2n+2} \left(w^2 + \frac{a^2}{b^2} \sec^{2n^{-1}}\theta\right)^{-n-3/2} \\ &= \frac{\rho_0}{m\pi^{1/2}} \frac{(n + \frac{1}{2})!}{n!} (2nh^2)^{n+1} (w^2 + 2nh^2 \sec^{2n^{-1}}\theta)^{-n-3/2}. \end{aligned} \tag{19}$$

In this form it will be seen that f is a function of the total energy only, for we have

$$\frac{d\Omega}{dz} = K(z) = -\left(\frac{C}{n}\right)^{1/2} ab^n \sin\theta,$$

and

$$\frac{dz}{d\theta} = (Cn)^{-1/2} ab^{-n-2} \cos^{-1-2n^{-1}}\theta.$$

Hence it follows that

$$\frac{d\Omega}{d\theta} = -\frac{a^2}{nb^2} \cos^{-1-2n^{-1}}\theta \sin\theta,$$

and

$$\Omega = -\frac{1}{2}\frac{a^2}{b^2}\cos^{-2n^{-1}}\theta = -nh^2\sec^{2n^{-1}}\theta.$$

In terms of Ω, (19) can be written in the form

$$f(z,w) = \frac{\rho_0}{m\pi^{1/2}}\frac{(n+\frac{1}{2})!}{n!}(2nh^2)^{n+1}(w^2-2\Omega)^{-n-3/2}.$$

This shows that, as expected, f is really a function of a single variable, $(\frac{1}{2}w^2-\Omega)$, the total energy.

From (17) and (19) it can also be shown that the "isothermal" solution, considered in the previous section, is the limiting case as n tends to infinity of the solution just obtained. From (17) we have

$$\begin{aligned} z &= \frac{1}{(Cn)^{1/2}}ab^{-n-2}\int_0^\phi \sec^{1+2n^{-1}}\phi\,d\phi \\ &= \left(\frac{h^2}{2\pi G\rho_0}\right)^{1/2}\int_0^\phi \sec^{1+2n^{-1}}\phi\,d\phi. \end{aligned}$$

As n tends to infinity, we have

$$\begin{aligned} z &\to \left(\frac{h^2}{2\pi G\rho_0}\right)^{1/2}\int_0^\phi \sec\phi\,d\phi \\ &= \left(\frac{h^2}{2\pi G\rho_0}\right)^{1/2}\log(\sec\theta+\tan\theta). \end{aligned}$$

Hence it follows that

$$\sec\theta\to\cosh\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\},\quad \cos\theta\to\operatorname{sech}\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\},$$

$$\sin\theta\to\tanh\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}.$$

Now we also have

$$\begin{aligned} K &= -\left(\frac{C}{n}\right)^{1/2}ab^n\sin\theta \\ &= -2h(2\pi G\rho_0)^{1/2}\sin\theta \\ &\to -2h(2\pi G\rho_0)^{1/2}\tanh\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}. \end{aligned}$$

This is the value obtained in Section 3. Similarly we have for the velocity-distribution

$$\begin{aligned} f(z,w) &= \frac{\rho_0}{m\pi^{1/2}}\frac{(n+\frac{1}{2})!}{n!}(2nh^2)^{n+1}(w^2+2nh^2\sec^{2n^{-1}}\theta)^{-n-3/2} \\ &= \frac{\rho_0}{m\pi^{1/2}}\frac{(n+\frac{1}{2})!}{n!}(2nh^2)^{-1/2}\cos^{2+3n^{-1}}\theta\,(1+w^2/2nh^2\sec^{2n^{-1}}\theta)^{-n-3/2} \\ &= \frac{\rho_0}{(2\pi)^{1/2}mh}\frac{(n+\frac{1}{2})!}{n!\,n^{1/2}}\cos^{2+3n^{-1}}\theta\,(1+w^2/2nh^2\sec^{2n^{-1}}\theta)^{-n-3/2}. \end{aligned}$$

As $n\to\infty$, $(n+\frac{1}{2})!/n!\,n^{1/2}\to 1$, and $(1+w^2/2nh^2\sec^{2n^{-1}}\theta)^{-n-3/2}\to\exp(-w^2/2h^2)$. It therefore follows that

$$\begin{aligned} f(z,w) &\to \frac{\rho_0}{(2\pi)^{1/2}mh}\cos^2\theta\exp(-\tfrac{1}{2}w^2/h^2) \\ &= \frac{\rho_0}{(2\pi)^{1/2}mh}\operatorname{sech}^2\left\{\frac{z}{h}(2\pi G\rho_0)^{1/2}\right\}\exp(-\tfrac{1}{2}w^2/h^2). \end{aligned}$$

Thus the distribution given by (17) and (19) tends to the "isothermal" solution as the parameter n tends to infinity.

5. *A density distribution in which the thickness is finite.*—In Section 4 we confined attention to systems in which very large values of q could occur. Since there was no restriction on the kinetic energy, the system had necessarily to be infinite in extent. We now examine a case in which the system is finite, and in which, therefore, the value of q is bounded. Proceeding as in Section 4, we arrive at the equation for Q, namely,

$$\frac{1}{Q}\frac{dQ}{dq} = \frac{q}{lq^2+n}.$$

As before, the case in which the constant l is zero need not be considered, since it leads only to the "isothermal" solution. If $l \neq 0$, the general solution is that $Q = A(q^2+n/l)^{1/2l}$. If n/l is positive, this expression places no restriction whatever on the values of q. This is the case investigated in the previous section. If, however, n/l is negative, $-a^2$, say, then as q passes through either of the values $\pm a$, the quantity (q^2-a^2) changes sign. Since Q is essentially a positive number, this suggests how we can restrict the values of q. Let us define Q as follows:

$$Q = A(a^2-q^2)^{1/2l}, \quad |q| \leqslant a; \qquad Q = 0, \quad |q| > a.$$

(If we were to take $Q = A(q^2-a^2)^{1/2l}$, then $|q| \geqslant a$, and zero velocities would never occur. This cannot correspond to a physical system.) The index l must now be positive, since Q must be finite for $|q| \leqslant a$. We also have

$$1 = A\int_{-a}^{+a}(a^2-q^2)^{1/2l}\,dq, \qquad \overline{q^2} = A\int_{-a}^{+a} q^2(a^2-q^2)^{1/2l}\,dq.$$

From these equations we can show that $\overline{q^2} = a^2/(3+1/l)$. It is convenient to put $3+1/l = 2n$, where $n > \frac{3}{2}$.

The functions $K(z)$ and $s(z)$ now satisfy the differential equations

$$l\frac{d}{dz}\left\{\log\left(s\frac{dK}{dz}\right)\right\} = -\frac{d}{dz}(\log s),$$

$$a^2\frac{ds}{dz} = -s^3K.$$

The first can be integrated, to give

$$\frac{dK}{dz} = -Cs^{-2(n-1)}. \tag{20}$$

Combining the second equation with this, we obtain

$$K\frac{dK}{dz} = Ca^2s^{-(2n+1)}\frac{ds}{dz}.$$

On integration, this becomes

$$K^2 = -\frac{Ca^2}{n}(s^{-2n}-b^{-2n}),$$

where b is the value of $s(z)$ in the plane $z=0$, in which K vanishes. Clearly b is the minimum value attained by $s(z)$, so that it is convenient to write

$$s^{2n} = b^{2n}\sec^2\theta,$$

or

$$s = b\sec^{n-1}\theta,$$

where θ is a function of z vanishing at $z=0$. The expression for K now takes the form

$$K = -\left(\frac{C}{n}\right)^{1/2} ab^{-n} \sin\theta.$$

It follows that

$$\frac{dK}{dz} = -\left(\frac{C}{n}\right)^{1/2} ab^{-n} \cos\theta \frac{d\theta}{dz}.$$

But from (20) we have

$$\frac{dK}{dz} = -Cb^{-2n+2} \cos^{2-2n^{-1}}\theta.$$

Thus z is connected with the parameter θ by the integral

$$z = \frac{1}{(Cn)^{1/2}} ab^{n-2} \int_0^\theta \sec^{1-2n^{-1}}\phi \, d\phi.$$

The index of $\sec\phi$ is less than unity, so that the integral is finite as θ tends to $\pi/2$.

The density is given by the formula

$$\rho = -\frac{1}{4\pi G}\frac{dK}{dz}$$

$$= \frac{C}{4\pi G} b^{-2n+2} \cos^{2-2n^{-1}}\theta.$$

Since $n > \frac{3}{2}$, the power of $\cos\theta$ is positive, so that the density has a maximum value $\rho_0 = Cb^{-2n+2}/4\pi G$ in the plane $z=0$.

It is convenient to introduce a second parameter h, where h^2 is the mean-square velocity at the level of maximum density. We have obtained the equation

$$\overline{q^2} = a^2/2n.$$

The mean-square velocity at the level z is therefore given by the expression

$$\overline{w^2} = \overline{q^2}/s^2 = a^2/2ns^2.$$

In the plane $z=0$ this simply becomes

$$h^2 = a^2/2nb^2.$$

In terms of ρ_0 and h, we obtain the following relations between $K(z)$, the force, $\rho(z)$, the density, z, the displacement, and the variable θ:—

$$K = -2h(2\pi G\rho_0)^{1/2} \sin\theta,$$

$$\rho = \rho_0 \cos^{2-2n^{-1}}\theta,$$

$$z = \frac{h}{(2\pi G\rho_0)^{1/2}} \int_0^\theta \sec^{1-2n^{-1}}\phi \, d\phi.$$

The total mass in a unit column, M, is given by

$$M = 2h(\rho_0/2\pi G)^{1/2} \int_0^{\pi/2} \cos^{2-2n^{-1}}\theta \sec^{1-2n^{-1}}\theta \, d\theta$$

$$= (2\rho_0 h^2/\pi G)^{1/2}.$$

All three solutions, therefore, give the same relation between M, ρ_0 and h.

The total thickness of the distribution is finite, and the actual value is

$$\frac{2h}{(2\pi G\rho_0)^{1/2}} \int_0^{\pi/2} \sec^{1-2n^{-1}}\phi \, d\phi = \frac{h}{(2\pi G\rho_0)^{1/2}} \frac{(n^{-1}-1)!(-\frac{1}{2})!}{(n^{-1}-\frac{1}{2})!}.$$

This tends to infinity with n.

The velocity-distribution function can also be expressed in terms of ρ_0 and h, and we then have

$$f=\frac{\rho_0}{m\pi^{1/2}}\frac{(n-1)!}{(n-\frac{1}{2})!}(2nh^2)^{1-n}(2nh^2\cos^{2n^{-1}}\theta-w^2)^{n-3/2}.$$

Here again we have the result that as n tends to infinity the distribution tends exactly to the "isothermal" case. The distribution has another remarkable feature, for the mass-density is distributed in the same way as a gas would be, if it obeyed an adiabatic law, was stratified in plane parallel layers, and was in equilibrium under its own gravitational attraction. The index n in the above solution and the ratio, γ, of the specific heats of the gas, must be related by the equation

$$\frac{1}{\gamma}=1-\frac{1}{n},$$

or

$$\gamma=1+1/(n-1).$$

Hence, in terms of the polytropic index, p, of the gas,

$$n=p+1.$$

The solution also required that $n>\frac{3}{2}$, which implies $p>\frac{1}{2}$, or $\gamma<3$.

The method of solution which yielded the results given in the last two sections has also been applied to the problem of non-steady distributions. However, it was found that if we assumed a velocity-distribution of the form

$$f(t,z,w)=Z(t,z)Q\{ws(t,z)\},$$

then we must have either $s=$constant—which leads to the artificially non-steady solution quoted at the end of Section 3—or else both Z and s must be independent of t. The second alternative implies that there can only be steady solutions of the desired form, and the details of these solutions have already been given.

6. *The comparison of the three models with observation.*—It is convenient to denote by I, II and III the models obtained in Sections 3, 4 and 5. Then we have seen that I is an exceptional case, separating models of type II from those of type III. It is infinite in extent, and corresponds to the density-distribution of an isothermal gas in equilibrium in its own gravitational field. Model II is also infinite in extent, and corresponds mathematically to the equilibrium configuration of a gas for which γ, the ratio of the specific heats, is less than unity. Model III has a finite thickness, and corresponds to the gas obeying the adiabatic law, with $1<\gamma<3$. The close similarity between the three is evident from Table I, in which the chief physical quantities are expressed as functions of a parameter.

TABLE I

The density, gravitational force, and distance from the central plane expressed in terms of a parameter, θ

Model	I	II	III
ρ/ρ_0	$\cos^2\theta$	$\cos^{2+2n^{-1}}\theta$	$\cos^{2-2n^{-1}}\theta$
$-K/(8\pi G\rho_0h^2)^{1/2}$	$\sin\theta$	$\sin\theta$	$\sin\theta$
$z(2\pi G\rho_0/h^2)^{1/2}$	$\int_0^\theta \sec\phi\,d\phi$	$\int_0^\theta \sec^{1+2n^{-1}}\phi\,d\phi$	$\int_0^\theta \sec^{1-2n^{-1}}\phi\,d\phi$

In the table ρ_0 is the density in the central plane of the system, and h is the root-mean-square velocity in that plane. The quantity z is the height above the plane, and K is the resultant gravitational force, measured positively in the direction of z increasing. As we have seen, ρ and K can be expressed explicitly as functions of z for solution I, but this parametric representation is more convenient when we wish to compare the three models. It is evident from the table that M, the total mass in a column of unit cross-section, is the same in all three cases, namely $\{2\rho_0 h^2/\pi G\}^{1/2}$.

The problem of a system of gravitating stars in which the surfaces of equal density are parallel planes has an application to the distribution of velocity and of density in the galactic system, normal to the galactic plane. The system is extremely flattened, and though there is a considerable central condensation, it appears that in the region of the Sun the surfaces of constant density are effectively parallel, at least up to 500 parsecs on either side of the galactic plane. Parallel to that plane, the dominant forces are the attraction of the central condensation and the centrifugal force arising from galactic rotation. The equality of these two leads to an estimate of the mass of the central nucleus. The motions perpendicular to the plane are controlled largely by the attraction of the parallel layers of material in the immediate vicinity. The attraction of the central nucleus has a component perpendicular to the galactic plane; this is comparatively small up to distances of 500 parsecs from the plane, though ultimately, of course, it is the dominant force. However, the effect of this component is ignored in the present paper.

In making a comparison with the galactic system in the region of the Sun, we are able to use the observational data collected by Oort (**7**), who has considered the same problem. For the density in the galactic plane (the quantity ρ_0), Oort derives the value 0·092 solar masses per cubic parsec. For our purpose it is unfortunate that we have to adopt this value, since Oort obtained it almost by reversing the steps we shall take. His procedure was to estimate the variation of density with z for different spectral types, and hence to calculate the variation of gravitational force required to produce the observed densities. He then considered what combination of material in ellipsoids, together with a massive central nucleus, would yield this distribution of force, and in this way estimated the local mass-density. However, the value for ρ_0 of 0·1 solar masses per cubic parsec, which we adopt, is certainly of the right order of magnitude. More than one-third of this figure can be accounted for by stars of absolute magnitude brighter than +13·5, and there are also intrinsically faint stars and interstellar matter to be added.

For the value of the root-mean-square velocity in the plane of maximum density, i.e. the galactic plane, we take $h = 20$ km./sec. This value is certainly too large for the B-type stars, probably too large for the A type, but is about correct for the F, G, K and M types which are more numerous. Oort assumed that the distribution of velocities was Gaussian, and found that it was usually necessary to add two such distributions to reproduce the observations. The dominant Gaussian distribution usually had a root-mean-square velocity of the order of 20 km./sec.

With these two figures, the value of M, the total mass in a column of unit cross-section, comes out to be 77 $M_\odot$ per square parsec. From his determination of density gradients, Oort concludes that if the proportion of dark matter is the

same at all levels, then M is 80 solar masses, whereas if the dark matter is all within 200 parsecs of the galactic plane, then the total mass in such a column would be only 51 solar masses. Our value of $77\,M_\odot$ is therefore in general agreement with Oort's estimate. The order of magnitude of this result can be checked in a different way. The total mass inside a column of which the cross-section is a circle of radius 15,000 parsecs comes out to be $5{\cdot}4 \times 10^{10}\,M_\odot$. The galactic system has dimensions of this order; from studies of galactic rotation Oort (**9**) finds that about $1{\cdot}3 \times 10^{11}\,M_\odot$ are concentrated in the nucleus, and 6×10^{10} solar masses are spread fairly uniformly in a disk. The mass in the disk, which must be mainly responsible for accelerations perpendicular to the galactic plane, is entirely comparable with the value we have obtained.

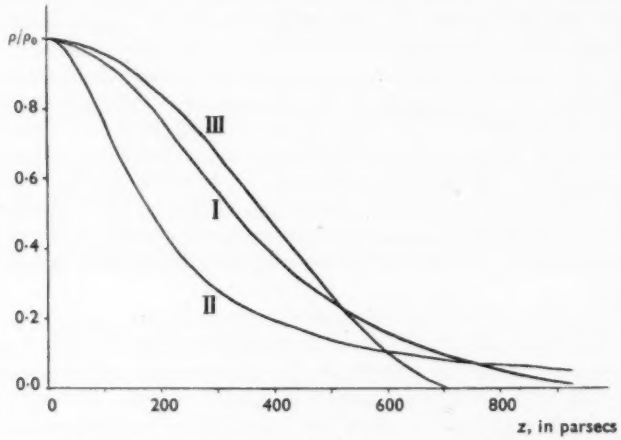


FIG. 1.—*The variation of density with distance from the galactic plane.*

The variation of density with distance from the galactic plane is illustrated in Fig. 1. Three curves are given, corresponding to the three different models. In the cases of II and III the index n is at our disposal. The variation of n will not affect the general nature of the curves, but in either case, as n increases, the curve will approach curve I. The value of n selected for curve II is $\frac{2}{7}$; this is particularly convenient, since the integral for z can then be evaluated in simple terms. This is the case whenever n is of the form $2/(2p+1)$, where p is an integer. For curve III the index n was taken to be $\frac{5}{2}$, which corresponds to a value, $\gamma = \frac{5}{3}$, for the adiabatic index. For this curve the density falls to zero at 708 parsecs. The area under all the curves is, of course, the same, namely $M/2\rho_0$. Curve II would appear to give the best agreement with observation. Oort estimates that in the galactic system ρ/ρ_0 falls to 0·1 at a height of 500 parsecs; the value at this level, as given by solution II, is 0·13, and this must be regarded as quite satisfactory.

A more detailed comparison with observed densities will now be given. Oort has evaluated the variation of density normal to the galactic plane in the following way. Stars which are within 50° of either galactic pole are first grouped according to absolute magnitude. Each group is then subdivided according to parallax. For each subdivision he then finds $\nu(z)$, the number of objects it contains, and z, the mean distance in parsecs from the galactic plane without regard to sign.

Finally he tabulates $\{\log \nu(z) - \log \nu(44)\}$ against z for each group. The two magnitude groups, $-0{\cdot}5$ to $+1{\cdot}5$, and $+1{\cdot}5$ to $+3{\cdot}5$, give the most reliable results, and these are reproduced in the second and third columns of Table II. The two columns are very consistent. In the next three columns are given the quantities, $\log\{\rho(z)/\rho(44)\}$, as calculated from solutions I, II and III. The models I and III differ very obviously from the observations, but there is extraordinarily good agreement between the values obtained from solution II and the observed quantities. Of course, the comparison of logarithms will tend to smooth out some of the discrepancies, but the agreement is much closer than one would have expected.

TABLE II

The variation of density with distance from the galactic plane, as observed, and as calculated

Distance, in parsecs, from the galactic plane, z	Logarithms of observed numbers of stars, $\log \nu(z) - \log \nu(44)$		Logarithms of calculated densities, $\log \rho(z) - \log \rho(44)$		
	$-0{\cdot}5 < M < +1{\cdot}5$	$+1{\cdot}5 < M < +3{\cdot}5$	I	II	III
44	0·000	0·000	0·000	0·000	0·000
69	−0·023	−0·036	−0·010	−0·032	−0·006
110	−0·104	−0·109	−0·025	−0·108	−0·018
174	−0·246	−0·251	−0·076	−0·248	−0·045
275	−0·465	−0·422	−0·198	−0·465	−0·135
438	−0·706	−0·738	−0·463	−0·748	−0·404
694	−1·113	...	...	−1·048	...

Another comparison which can be made with Oort's investigation is that of gravitational force. In Fig. 2 are shown three typical curves which give the variation of $(-K)$, the acceleration towards the central plane, as a function of z, the distance from the plane. The ordinate at each point is proportional to the area beneath the corresponding curve of Fig. 1. There is a maximum acceleration of $6{\cdot}75 \times 10^{-9}$ cm./sec.^{2}, which is attained by curve III at $z = 708$ parsecs, the level at which ρ vanishes. Curves I and II approach this limit asymptotically, as z tends to infinity.

The quantity K has also been calculated by Oort directly from the observations. If equation (2) is multiplied by w, and then integrated over all values of w, we obtain the equation

$$K(z) = \frac{1}{\nu} \frac{d}{dz} (\nu \overline{w^2}).$$

For objects of different spectral type the quantities $\overline{w^2}$ and $\nu(z)$ have been found, and Oort has obtained values of $K(z)$ for each type, using the above formula in the form

$$K(z) = \overline{w^2} \frac{d}{dz} (\log \nu). \qquad (21)$$

After applying corrections and weighting the determinations, Oort gives values of $K(z)$ for eight different values of z. These are the points marked in Fig. 2. In view of the close agreement we have found in Table II between Oort's $\log \nu(z)$ and our $\log \rho(z)$, we might have expected curve II to fit the points more closely.

The difference must be accounted for by the variation of $\overline{w^2}$ with z. The exact formula, which gives the curves in Fig. 2, is

$$K(z)=\overline{w^2}\frac{d}{dz}(\log \nu)+\frac{d}{dz}(\overline{w^2}). \qquad (22)$$

The points have been determined from (21) where $\overline{w^2}$ is a constant, namely, its value in the plane $z=0$. Since $(d/dz)(\log \nu)$ is negative, and since in model II $\overline{w^2}$ increases with z, the first term on the right-hand side of (22) is more negative than the corresponding term of (21). The other term in (22) is positive, and tends to equalize the two values of $K(z)$, though differences will remain.

There is one peculiarity of Oort's values of $K(z)$ which is quite contrary to the behaviour of $K(z)$ for a self-gravitating system. Irrespective of the *solution* of the differential equation, we always have for a stratified self-gravitating system the relation

$$\frac{dK}{dz}=-4\pi G\rho.$$

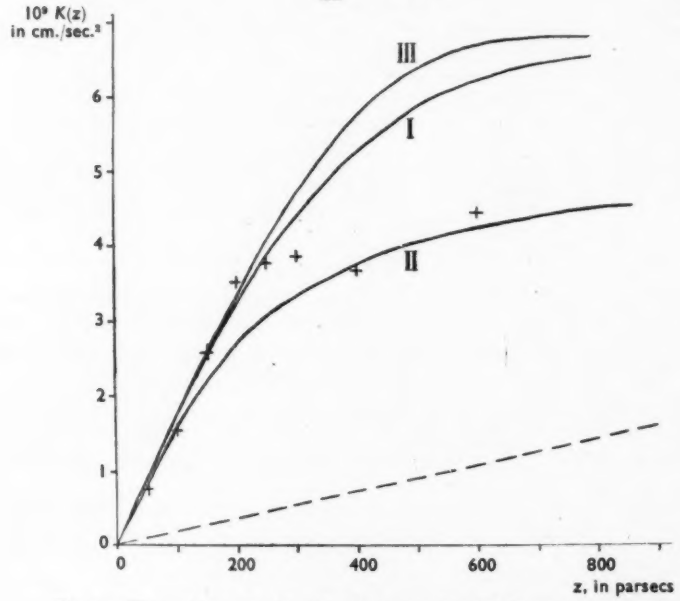


FIG. 2.—*The variation of acceleration due to gravity with distance from the galactic plane.*

Thus the slope for the curve of K as a function of z cannot change sign for a physical system. Such a change of sign, however, is certainly suggested by the values Oort has given. This effect may also be attributed to the variation of $\overline{w^2}$ with z, for which he has not allowed.

On Fig. 2 is also shown a broken line, which gives approximately the component towards the galactic plane of the acceleration produced by a central nucleus of mass $1{\cdot}3\times10^{11}\,M_\odot$, situated in the plane, but at a distance of 10,000 parsecs. For z less than 1000 parsecs this component is practically proportional to z and reaches a value of $1{\cdot}82\times10^{-9}$ cm./sec.^{2} at $z=1000$ parsecs. In comparing our theoretical results with the observed features of the galactic system, we have ignored

completely the influence of the central nucleus. From Fig. 2 it is obvious that the acceleration it produces is small, within the range of values of z considered, compared with that produced by the plane layers of matter. However, it is also clear that the effect of the nucleus is certainly not negligible.*

A final comparison with observation can be made, namely, in the distribution of velocities normal to the galactic plane. Oort has tabulated the distribution of w for stars of spectral types F, G, K and M within 20 parsecs of the Sun, grouping according to spectral type. In fact he subdivided the spectral classes further, but though there may be genuine differences in velocity-distribution in these subdivisions, the number of objects in each becomes so small that it is scarcely possible to recognize even the kind of distribution. In these circumstances we have retained Oort's three original groups only. The corresponding distributions are plotted in Fig. 3. The ordinate is the percentage of stars with velocities in a

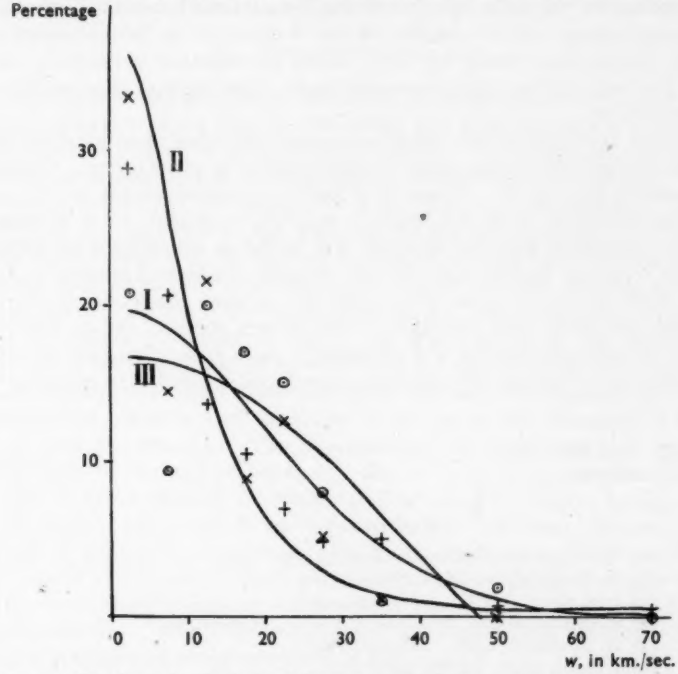


FIG. 3.—*The observed and calculated distributions of velocity components, w, in the region of the Sun.*

range of 5 km./sec., centred on the corresponding abscissa. The distributions for F type, G type, and the K and M types together are indicated by the symbols

* In a further paper on this subject Oort and van Woerkom (10) have used the density distribution of the RR Lyrae variables to determine $K(z)$ in the range from 1000 to 10,000 parsecs. This is far beyond the range to which the present method will apply, but it is of interest to see that $K(z)$ is constant and equal to $8{\cdot}4 \times 10^{-9}$ cm./sec.^{2}. This is by no means inconsistent with our present results, for Oort and van Woerkom attribute the greater part of $K(z)$ in their range to the effect of the central condensation. The remaining part, due to the extended distribution of matter, falls short of our value of $6{\cdot}75 \times 10^{-9}$, but beyond 2000 parsecs this is to be expected, for we have been using a model extending to infinity in all directions in the galactic plane. Such a model should give good results up to 500 parsecs, and will be quite reasonable up to 1000 parsecs.

$\times$, $\odot$ and $+$ respectively. The three curves give the distribution of velocities in the region of the Sun, as predicted by solutions I, II and III. Clearly the scatter of points is so wide that no curve will fit them, but the general run of the points would be fairly consistent with a distribution lying between curves I and II. Such a curve could be obtained by taking a solution of the type II, but with a value of n greater than $\frac{2}{7}$.

If a similar analysis were undertaken for stars in a larger volume of space, some of the irregularities in the velocity-distribution would be smoothed out. But Oort has pointed out that even in the sphere of radius 20 parsecs there may be a deficiency in the observed number of stars with very small velocities. Such a deficiency has a two-fold effect on the observed velocity-distribution; it lowers the ordinates for small values of w, and it leads to too large a value of the mean-square velocity. However, even if we knew neither the total number of stars in the group nor the mean-square velocity, it would still be possible—though perhaps rather tedious—to fit a solution of one of the types we have obtained to an observed distribution, which was only known for values of w between, say, 15 and 50 km./sec. At any rate this would seem to be one way of obtaining a reliable comparison.

The method developed in this paper has enabled us to find three solutions for the problem of a self-gravitating stellar system stratified in plane layers. Clearly it can be used to find others. However, it leaves undetermined the stability of the solutions obtained. This will require separate treatment, since it seems necessary to investigate both the stability of a model in which $f(t, z, w)$ differs slightly from a known solution, and also the stability of a known solution under the action of the stochastic forces which arise during close encounters. These, however, are essentially new problems. The present method can be applied equally well to spherical systems, and probably to more general systems, though the great advantage of having only one space coordinate will be lost in the general case.

The University, Manchester, 13:
1949 *November* 21.

References

(1) W. M. Smart, *Stellar Dynamics*, p. 319, 1938.
(2) A. S. Eddington, *M.N.*, **76**, 37, 1915.
(3) J. H. Oort, *B.A.N.*, **4**, No. 159, 1928.
(4) S. Chandrasekhar, *Principles of Stellar Dynamics*, 1942.
(5) G. L. Camm, *M.N.*, **101**, 431, 1941.
(6) J. Jeans, *M.N.*, **76**, 78, 1915.
(7) J. H. Oort, *B.A.N.*, **6**, No. 249, 1932.
(8) S. Chandrasekhar, *loc. cit.*, p. 190.
(9) J. H. Oort, *B.A.N.*, **3**, No. 120, 1927.
(10) J. H. Oort and van Woerkom, *B.A.N.*, **9**, No. 338, 1941.

ON THE INSTABILITY OF SMALL PLANETARY CORES (I)

W. H. Ramsey

(Received 1950 April 13)

Summary

The instability of small planetary cores, which was encountered in one of the author's previous papers (*M.N.*, **108**, 406, 1948), is investigated more fully. It is pointed out that the phenomenon may have a bearing on the origin of meteorites and asteroids.

---

1. In previous papers (**1**, **2**) the author has suggested that the jump in density at the boundary of the Earth's core is due to a phase transition under pressure, and not to a change in chemical composition. This suggestion, together with the assumptions current in geophysics, implies that the variation in density within the Earth is due primarily to the influence of pressure. Seismology provides the pressure-density relationship for the material of the Earth (**3**). The mean densities of the other terrestrial planets can be computed on the basis of this pressure-density relationship, assuming that these planets have the same chemical composition as the Earth; this has been carried out in another paper (**1**). An interesting phenomenon which was encountered in the course of these calculations is that very small cores are unstable and so would not occur in nature. This feature is a direct consequence of the main assumption underlying the calculations, namely, that the jump in density at the boundary of the core is controlled by pressure. The phenomenon was explained very briefly in an Appendix to the earlier paper (**1**), but a more comprehensive treatment is desirable for a number of reasons. First, several people, notably Professor Jeffreys (**4**), have shown an interest in the problem. Second, the same phenomenon will be encountered again in a forthcoming paper on the internal structures and constitutions of the major planets. Third, it will be shown that there is a possibility that the phenomenon has a bearing on the origin of meteorites and asteroids. The phenomenon and its implications will be examined in considerable detail in the two parts of this paper.*

2. In its most general form the problem is mathematically cumbersome. The phenomenon can be understood most easily by treating a simple example which can be handled algebraically. It will therefore be assumed in Part I of this paper that the planets are composed of a material which has a constant density ρ_0 at all pressures below a critical pressure p_c, and a constant density $\lambda\rho_0$, which is greater than ρ_0, at all higher pressures. In other words, it will be assumed that the material is incompressible, both above and below the critical pressure. In Part II, Lighthill will show that the main conclusions of this part of the paper are valid also in the general case of a compressible material.

* For Part II, see M. J. Lighthill, *M.N.*, **110**, 339, 1950.

Let M be the mass of the planet under consideration, R the total radius, R_c the radius of the core, and P the central pressure. The pressure at the boundary of the core must be equal to the critical pressure p_c; that is,

$$p_c=(2\pi/3)G\rho_0^2\{R^2-R_c^2+2(\lambda-1)R_c^3(R_c^{-1}-R^{-1})\}, \tag{1}$$

where G is the constant of gravitation. It is convenient to introduce the mass M_0 and the radius R_0 of the planet whose central pressure is equal to the critical pressure p_c:

$$p_c=(2\pi/3)G\rho_0^2R_0^2, \qquad M_0=(4\pi/3)\rho_0R_0^3. \tag{2}$$

The condition (1) that the boundary of the core is at the critical pressure p_c can then be written in the simpler form

$$R^2-2(\lambda-1)R_c^3R^{-1}+\{(2\lambda-3)R_c^2-R_0^2\}=0. \tag{3}$$

When there is a core the central pressure P is given by

$$P/p_c=1+\lambda^2(R_c/R_0)^2, \tag{4}$$

and the total mass M is

$$M=(4\pi/3)\rho_0\{R^3+(\lambda-1)R_c^3\}. \tag{5}$$

The three equations (3), (4) and (5) contain all the physical assumptions made. The problem is to find the combinations of the variables M, R, R_c and P which are compatible with these equations.

3. To begin with, it will be proved that the central pressure P always determines the planetary configuration uniquely. For this reason the central pressure will be made the independent variable in the problem. First, suppose that P is less than the critical pressure p_c, so that there is no core. The mass M and the radius R are given unambiguously by the equations

$$M/M_0=(P/p_c)^{3/2}, \quad R/R_0=(P/p_c)^{1/2}. \tag{6}$$

The mass M and the radius R increase monotonically with the central pressure P, attaining the values M_0 and R_0 when P is equal to p_c. Second, suppose that P is greater than the critical pressure p_c, so that there is a core. The radius R_c of the core is uniquely determined by P through equation (4). Also, the radius of the core determines the total radius R uniquely. This follows because the left-hand side of equation (3), for any given value of R_c, is an increasing function of R; and it is negative when R is put equal to R_c, and it increases as R^2 for large values of R. Hence, to each value of R_c there corresponds one, and only one, value of R which satisfies equation (3); and this value of R is necessarily greater than R_c. When both R and R_c are known, the mass M is given unambiguously by expression (5). Hence, in all cases, the central pressure P uniquely determines all the properties of the planet. This result is important, as it will be shown below that several configurations can have the same total radius R, and also that several configurations can have the same total mass M.

4. Consider first the behaviour of the total radius R as a function of the central pressure P. When P is less than the critical pressure p_c, it follows from equation (6) that R is an increasing function of P, so that P and R uniquely determine one another. But R is not always an increasing function of P when the central pressure exceeds the critical pressure. This is evident from equation (3); for the total radius R has the value R_0 when R_c is zero, and also when

$$R_c/R_0=(2\lambda-3)/(2\lambda-2), \tag{7}$$

and this second possibility is physically admissible if λ is not less than $\frac{3}{2}$. The derivative of R with respect to R_c can readily be derived from equation (3), and it is

$$dR/dR_c = (4\pi/3)\rho_0 R R_c M^{-1}\{3(\lambda-1)R_c-(2\lambda-3)R\}, \tag{8}$$

where M is the total mass. The derivatives (dR/dP) and (dR/dR_c) invariably have the same sign, as it follows from equation (4) that (dP/dR_c) is always positive:

$$dP/dR_c = 2\lambda^2 p_c R_c/R_0^2. \tag{9}$$

As has been stated earlier, λ is assumed to be greater than unity as the high pressure modification will be the denser. The derivative (8) is necessarily positive if λ is less than $\frac{3}{2}$; R is therefore an increasing function of R_c, and so also of the central pressure P. Hence, if λ is less than $\frac{3}{2}$, R is an increasing function of P, both above and below the critical pressure. Any given value of R therefore uniquely determines the central pressure P, and so all the properties of the planet. On the other hand, the derivative (8) may be negative if λ exceeds $\frac{3}{2}$, and indeed it always is negative for sufficiently small values of R_c. In this case $(\lambda>\frac{3}{2})$ the derivatives (dR/dR_c) and (dR/dP) are negative for ratios (R_c/R) smaller than

$$R_c/R = (2\lambda-3)/(3\lambda-3), \tag{10}$$

and they are positive for all larger values of (R_c/R). It may easily be shown from equation (8) that the ratio (R_c/R) is always an increasing function of R_c, and so also of P. When λ exceeds $\frac{3}{2}$ the behaviour of R as a function of P is therefore as follows. (Diagrammatically the behaviour of R is similar to the behaviour of M as a function of P, and the latter is shown in Fig. 1.) As P is increased from zero to the critical pressure p_c, R increases uniformly from zero to the value R_0 (*cf.* equation (6)). When the central pressure P is increased beyond the critical pressure, the total radius R steadily decreases until it reaches a minimum value R', when equation (10) is satisfied. Thereafter it increases continuously with P, and it again attains the value R_0 when equation (7) is satisfied. It follows that one, and only one, central pressure is compatible with a given value of the total radius R, provided that the given value of R does not lie in the range

$$R' \leqslant R \leqslant R_0. \tag{11}$$

This means that the planet can be specified unambiguously by its total radius R, provided that R does not lie in the range (11). On the other hand, for each value of R within the range (11) there are three possible values of P, two being greater than and one less than the critical pressure p_c. This means that a given total radius R in the range (11) is compatible with three different planetary configurations, two with cores and one without a core. In general, the three configurations have different total masses.

A complicated explicit expression for R in terms of R_c can, if required, be derived from equation (3) by the standard procedure for solving a cubic equation. But, for small values of R_c, the behaviour of R is given more simply by a power series development in (R_c/R_0):

$$R/R_0 = 1-(\lambda-\tfrac{3}{2})(R_c/R_0)^2+(\lambda-1)(R_c/R_0)^3+\ldots. \tag{12}$$

When λ is greater than $\frac{3}{2}$ this series exhibits the minimum described above. The region of the minimum can always be investigated by this method; the absolute convergence of series (12) in ensured because, as is evident from equation (7), the ratio (R_c/R_0) is always less than unity in the neighbourhood of the minimum. For

very massive planets, R_c and R are nearly equal. This is evident from the expansion in descending powers

$$R/R_c = 1 + \tfrac{1}{2}\lambda^{-1}(R_0/R_c)^2 + \ldots, \tag{13}$$

which is valid if R_c is large compared with R_0. The thickness $(R - R_c)$ of the mantle decreases rapidly as the size of the planet increases. This is because the boundary of the core is at a fixed pressure, and gravity is larger in more massive planets.

5. Consider next the behaviour of the total mass M as a function of the central pressure P. When P is less than the critical pressure p_c it follows from equation (6) that M continually increases with P and attains the value M_0 when P is equal to p_c. There is a core when P exceeds the critical pressure, and the total mass is given by equation (5). According to equation (9) the derivative (dP/dR_c) is always positive, so that (dM/dP) and (dM/dR_c) invariably have the same sign. The latter derivative is

$$dM/dR_c = 4\pi\rho_0\{R^2(dR/dR_c) + (\lambda - 1)R_c^2\}, \tag{14}$$

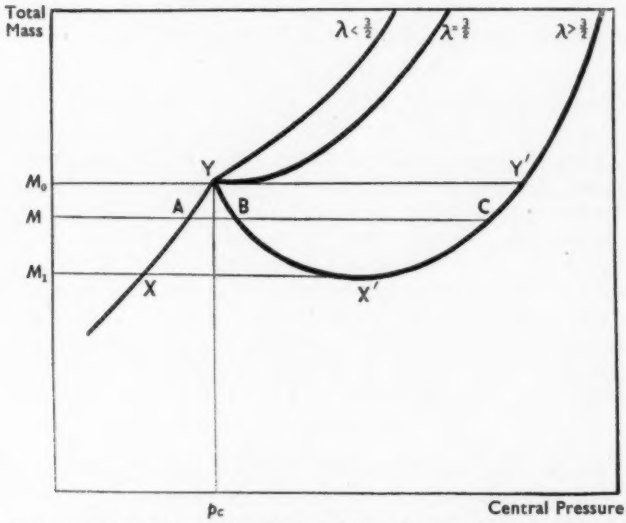


FIG. 1.—*The behaviour of the total mass M as a function of the central pressure P.*

which can be negative only if (dR/dR_c) is negative. But it has been shown above that (dR/dR_c) is always positive if λ is less than $\frac{3}{2}$. Hence, for λ less than $\frac{3}{2}$, the total mass M is an increasing function of P, both above and below the critical pressure; M therefore uniquely determines the central pressure P, and so all the properties of the planet. On the other hand, it has been shown above that (dR/dR_c) can be negative if λ exceeds $\frac{3}{2}$. It will now be shown that (dM/dR_c) can also be negative if this condition is fulfilled. It follows from equations (8) and (14) that

$$dM/dR_c = 4\pi\rho_0(M_0/M)(R/R_0)^3 R R_c \mathcal{F}, \tag{15}$$

where

$$\mathcal{F} = -(2\lambda - 3) + 4(\lambda - 1)(R_c/R) + (\lambda - 1)^2(R_c/R)^4. \tag{16}$$

Clearly (dM/dR_c) is negative if, and only if, the factor $\mathcal{F}$ is negative, and this is possible only if λ exceeds $\frac{3}{2}$. But if λ does exceed $\frac{3}{2}$, $\mathcal{F}$ is always negative for sufficiently small values of the ratio (R_c/R); consequently (dM/dR_c) and (dM/dP)

are also negative. $\mathscr{F}$ is an increasing function of (R_c/R), and it is negative when (R_c/R) is zero and it is positive when (R_c/R) is unity. There is therefore one, and only one, value of (R_c/R) which makes $\mathscr{F}$ equal to zero. $\mathscr{F}$ is negative for all smaller values of (R_c/R), and it is positive for all larger values. It has been pointed out earlier that (R_c/R) is always an increasing function of the central pressure P. Hence the value of (R_c/R) which makes $\mathscr{F}$ equal to zero uniquely determines a central pressure P_1. The derivative (dM/dP) is negative if the central pressure P lies in the range

$$p_c < P < P_1, \tag{17}$$

and it is positive if P lies outside this range. For λ greater than $\frac{3}{2}$, the behaviour of M as a function of P is therefore as follows (*cf.* Fig. 1). As P is increased from zero to the critical pressure p_c, M continually increases to the value M_0 (*cf.* equation (6)). When the central pressure is increased beyond the critical pressure, M steadily decreases with P until P reaches the value P_1; here the mass M is a minimum*, say M_1. The mass M is an increasing function of P for all central pressures greater than P_1. It follows that one, and only one, central pressure is possible for a planet whose mass M does not lie in the range

$$M_1 \leqslant M \leqslant M_0. \tag{18}$$

This means that a planet whose mass does not lie in the range (18) has only one possible equilibrium configuration. In other words the mass determines unambiguously the total radius, mean density, etc. On the other hand, for each mass in the range (18) there are three possible values of P, one being less than and two greater than the critical pressure p_c. This means that a planet whose mass lies in the range (18) has three possible configurations, one without a core and two with cores; in Fig. 1 these configurations are represented by the points A, B and C. The order A, B, C is the order of increasing central pressure and also of decreasing total radius; A has no core, and B has a smaller core than C. Of the three configurations, C has the smallest gravitational potential energy. It was assumed in the calculations (1) on the terrestrial planets that only the configuration with the smallest energy would occur in nature. This led to the discontinuity in Fig. 1 of the previous paper (1), which illustrated the calculated variation of the mean density with the mass of the planet. The stabilities of the three configurations will be examined more closely below.

The region in which the total mass is a minimum can be investigated by the power series method which was used to investigate the minimum of the total radius. An expansion of the mass M in power of (R_c/R_0) can be obtained by inserting the expansion (12) for R into the expression (5) for the total mass

$$M/M_0 = 1 - 3(\lambda - \tfrac{3}{2})(R_c/R_0)^2 + 4(\lambda - 1)(R_c/R_0)^3 + \ldots, \tag{19}$$

which is the expression given in the Appendix to the previous paper (1). This series exhibits the minimum described above if λ is greater than $\frac{3}{2}$, and not otherwise. The series (19) is absolutely convergent when the mass is a minimum, as this always occurs at a smaller value of (R_c/R) than the minimum of the total radius. The total radius R is a minimum when (R_c/R) is given by equation (10); but it follows from equation (16) that (dM/dP) is positive at this value of (R_c/R), so that the mass M has already passed its minimum.

* The minimum value of R was denoted by R'. The minimum value of M will not be denoted by M' as the minimum mass and the minimum radius belong to different planetary configurations.

6. To illustrate the effect quantitatively, the constant λ will be put equal to 2. Table I gives a complete solution of the problem for central pressures in the neighbourhood of the critical pressure. The total radius R is a minimum when the central pressure P is $1{\cdot}43\,p_c$; the minimum value R' is only two per cent less than R_0. The total mass M is a minimum when the central pressure is $1{\cdot}24\,p_c$; the minimum mass M_1 is about three per cent less than M_0. No reasonable value of λ can lead to a large percentage difference between M_0 and M_1, or between R_0 and R'. For example, the density would have to increase by a factor of six at the critical pressure if R_0 and R' are to differ by 30 per cent. The density is unlikely to increase by this amount at a phase transition, even if one phase is metallic and the other non-metallic. The jump in density at the boundary of the Earth's core is not known precisely, but it probably corresponds to a value of λ slightly less than two (**3**). Using Bullen's (**3**) estimate of the density distribution within the Earth, it is found (**1**) by numerical integration that R_0 is about 6300 km. and M_0 is about 80 per cent of the mass of the Earth :

$$R_0 = 6300\text{ km.}, \quad M_0 = 0{\cdot}80\,M_E. \tag{20}$$

Adopting these values in Table I, it follows that the difference $(R_0 - R')$ is about 110 km., and the mass range $(M_0 - M_1)$ is about $0{\cdot}02\,M_E$. Numerical integration on the basis of Bullen's density distribution indicates a difference $(R_0 - R')$ somewhat larger than that given by Table I, and a mass range $(M_0 - M_1)$ somewhat smaller; the discrepancy arises from the assumption that the planets are composed of an incompressible material. But, for qualitative considerations and for rough numerical calculations, Table I gives a fair representation of the terrestrial planets. A planet whose mass lies in the critical range (18) can exist in configurations whose total radii differ by as much as 100 km. For example, the planet whose mass is M_0 can exist without a core and with a central pressure p_c, or with a core and with a central pressure $1{\cdot}55\,p_c$; it will be seen from Table I that the total radius of the former configuration is about 100 km. greater than that of the latter. If a planet were to switch over from one such configuration to another, the surface movements would be of the order of 100 km. The magnitude of such an upheaval is so great that it cannot be thought of in terms of earthquakes; the surface movement in an earthquake is, at most, a few metres.

TABLE I
($\lambda = 2$)

P/p_c	R_c/R_0	R/R_0	M/M_0
0·9	...	0·949	0·854
0·95	...	0·975	0·926
1·0	0·000	1·000	1·000
1·1	0·158	0·991	0·979
1·2	0·224	0·986	0·970
1·3	0·274	0·983	0 971
1·4	0·316	0·982	0·979
1·5	0·354	0·982	0·992
1·6	0·387	0·984	1·011
1·8	0·447	0·991	1·062
2·0	0·500	1·000	1·125

7. The total energy E of a planet will now be calculated as a function of its total mass M, its total radius R, and the radius R_c of its core. The ultimate aims of this

calculation are to investigate the stabilities of the three configurations represented by the points A, B and C in Fig. 1, and to estimate the energy which would be liberated if a planet were to undergo a transition from one such configuration to another. The zero of energy will be chosen to correspond to the material of the planet being dispersed at infinity and at zero pressure. The total energy E required to form the planet can be divided into two parts :

$$E = E_p + E_G. \tag{21}$$

This division corresponds to the fact that, in general, the placing of an element of material into its position in the planet involves two distinct processes. First, the physical conditions (in this case, the pressure) must be adapted to those prevailing in the planet; this gives rise to the term E_p in equation (21). Second, assembling the planet from the material dispersed at infinity alters the gravitational potential energy by an amount E_G, which is the second term in equation (21). As the material is assumed to be incompressible, the application of a pressure less than the critical pressure does not require energy. But at the critical pressure p_c the material collapses into the high pressure modification, and the work done per unit mass is

$$\epsilon_p = (\lambda - 1)p_c/\lambda\rho_0. \tag{22}$$

The energy required to convert sufficient material into the high pressure modification to form a planetary core of volume V_c is

$$E_p = (\lambda - 1)p_c V_c. \tag{23}$$

No energy is required to increase the pressure beyond p_c, as the high pressure modification also is assumed to be incompressible. Equation (23) therefore gives the first term in equation (21). The second term E_G can be evaluated by a standard procedure, and it can be written in the form:

$$E_G = -\tfrac{3}{5}(M^2G/R) - \tfrac{3}{5}(\lambda - 1)p_c V_c, \tag{24}$$

where G is the constant of gravitation. The total energy E is therefore

$$E = -\tfrac{3}{5}(M^2G/R) + \tfrac{3}{5}(\lambda - 1)p_c V_c, \tag{25}$$

where V_c is the volume of the core and is to be put equal to zero if there is no core.

8. The stability of a planetary configuration can be investigated in the following manner. Suppose that a small pressure π_0 is applied to the surface of the planet, so that the total radius changes by ΔR. The work done by the applied pressure is

$$\Delta E = -2\pi R^2 \pi_0 \Delta R. \tag{26}$$

The applied pressure π_0 can equally well be maintained by an added surface layer of thickness δ, where

$$\pi_0 = (MG/R^2)\rho_0\delta. \tag{27}$$

The mass of the added layer is

$$\delta M = 4\pi R^4 \pi_0 / MG. \tag{28}$$

But the addition of this layer leads to a planet of mass $(M + \delta M)$ in equilibrium, and its total radius is therefore $(R + \delta R)$, where

$$\delta R = (dR/dP)(dP/dM)\delta M, \tag{29}$$

where the derivatives (dR/dP) and (dM/dP) refer to planets in equilibrium and are given by equations (8), (9) and (15). But the radius of the planet must also be $(R+\Delta R+\delta)$; hence

$$\Delta R+\delta=(dR/dP)(dP/dM)\delta M. \tag{30}$$

Eliminating δ and δM with the aid of equations (27) and (28), equation (30) becomes

$$\pi_0=-\rho_0(MG/R^2)\{(dM/dP)-4\pi R^2\rho_0(dR/dP)\}^{-1}(dM/dP)\Delta R, \tag{31}$$

which gives the applied pressure π_0 in terms of ΔR, the change in the total radius. Inserting this expression for π_0, equation (26) becomes

$$\Delta E=2\pi MG\rho_0\{(dM/dP)-4\pi R^2\rho_0(dR/dP)\}^{-1}(dM/dP)(\Delta R)^2, \tag{32}$$

which is the energy required to change the total radius by ΔR. If there is a core, the factor in curly brackets in equation (32) is

$$(dM/dP)-4\pi R^2\rho_0(dR/dP)=4\pi R_c^2(\lambda-1)\rho_0(dR_c/dP), \tag{33}$$

which is always positive on account of equation (9). If there is no core this factor is exactly zero, as the material is assumed to be incompressible. But when the assumption of incompressibility is replaced by one of small compressibility, the factor concerned is small and positive. Hence the factor in curly brackets in equation (32) is always positive, so that ΔE invariably has the same sign as (dM/dP). The configuration is stable against small radial perturbations if ΔE is positive, and it is unstable if ΔE is negative. Hence the configuration is stable if (dM/dP) is positive, and it is unstable if (dM/dP) is negative. The configurations A and C in Fig. 1 are therefore stable, and B is unstable. This result is just what one would expect on the basis of Poincaré's principle of the exchange of stabilities (**5**). This principle asserts that if a system has one stable equilibrium state for certain values of a parameter λ, and if this state branches into three equilibrium states as λ is varied, then the two extreme states will be stable and the intermediate state will be unstable. This result has a simple geometrical interpretation in the energy-total radius diagram; a minimum can branch only into two minima separated by a maximum (*cf.* reference (**6**)). All configurations on the branch YX' in Fig. 1 are unstable, and these are the configurations with the smallest cores; for this reason the phenomenon is referred to as "the instability of small cores".

9. Fig. 2 shows schematically the dependence of the total energy E of a planet on its total radius R. The letters A, B, etc. refer to the same equilibrium configurations as in Fig. 1. The middle curve corresponds to a general mass M in the interval

$$M_1<M<M_0. \tag{34}$$

There are three equilibrium configurations A, B and C, of which A and C are stable and B is unstable. The energy of configuration A may be greater than that of C, or it may be less. A small disturbance could not cause a transition from A to C, or from C to A, because of energy barrier. A transition could occur during a collision with another planet. But a collision with another planet is very unlikely, even during the lifetime of a planet. Internal stresses, such as those developed by thermal expansion, are much too small to cause a transition. The central pressures of configurations A and C differ by a large fraction of the critical pressure, which is

about 1·4 million atmospheres for the change in phase at the boundary of the Earth's core. The strengths of the materials would not allow a deviation from hydrostatic conditions by anything approaching the order of magnitude of a million atmospheres. It would therefore appear that a planet can remain indefinitely in either configuration A or configuration C. The mass values M_0 and M_1 are special cases. The relationships between the total energy E and the total radius R for these masses are shown separately in Fig. 2; the top curve refers to the mass M_0, and the bottom curve to the mass M_1. For the mass M_0 the configurations A and B coalesce into Y, and for the mass M_1 the configurations B and C coalesce into X'. In both cases a maximum and a minimum in Fig. 2 coalesce to give a point of inflection. This is evident from equation (23); the derivative (dM/dP) vanishes for the configurations X' and Y, and so the second derivative (d^2E/dR^2) also vanishes. Y' is the only stable configuration of a planet of mass M_0, and X is the only stable configuration of a planet of mass M_1.

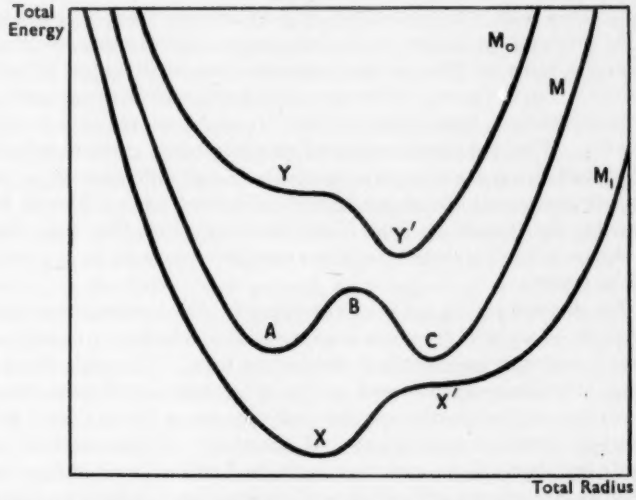


FIG. 2.—*The total energy E of a planet as a function of its total radius R. The top curve refers to a mass M_0, the bottom curve to a mass M_1, and the middle curve to a general mass in the range* (34).

10. Consider next the difference in energy between two configurations of the same planet. This can be evaluated directly from the general expression (25) for the total energy E of a planet in equilibrium. As an example, consider the configurations Y and Y' of the planet whose mass is M_0. The total energy of configuration Y is

$$E_0 = -\tfrac{3}{5}(M_0^2 G/R_0). \tag{35}$$

Let the energy of configuration Y' be $(E_0 + \Delta E_0)$. It follows from equation (25) that

$$\Delta E_0/E_0 = \{1-(\lambda-1)(R_c/R_0)^3\}^{-1/3} - \{1+\tfrac{1}{3}(\lambda-1)(R_c/R_0)^3\}, \tag{36}$$

where R_c is the radius of the core in configuration Y'. The ratio $(\Delta E_0/E_0)$ is very small for all reasonable values of λ. For λ equal to 2, the case treated in Table I,

$(\Delta E_0/E_0)$ is somewhat less than 0·1 per cent. But in absolute value ΔE_0 is very large. Using the constants (20) corresponding to the terrestrial planets, E_0 is about $1{\cdot}5 \times 10^{39}$ ergs; consequently the energy difference ΔE_0 is approximately

$$\Delta E_0 \simeq 10^{36} \text{ ergs.} \tag{37}$$

Earthquakes, on the average, release about 10^{20} ergs per year. At this rate the energy (37) is about equal to the total energy released by earthquakes during the whole of geological time. This gives some idea of the catastrophic consequences of a transition of a planet from one configuration to another. The energy released by an atomic bomb is about 10^{21} ergs. The energy difference between configurations X and X' of a planet of mass M_1 can be calculated in the same manner, and it is approximately equal to ΔE_0 above. The energy difference for the terrestrial planets will be greater than that given by equation (37) because of compression. The compression is greatest in the core, and this effectively increases the value of the constant λ. The energy difference (36) is rather sensitive to the value of λ. If λ is increased from 2 to 3, then $(\Delta E_0/E_0)$ is increased from 0·1 per cent to 3 per cent. As the jump in density at the boundary of the Earth's core is not known precisely, the value of ΔE_0 for the terrestrial planets is subject to considerable uncertainty. But the energy difference is probably sufficient to remove from the surface of the planet a mass between 0·001 M_E and 0·01 M_E, M_E being the mass of the Earth. The latent heat associated with the change in phase at the boundary of the core will also make a contribution to the energy difference ΔE_0. But neither the sign nor the magnitude of this latent heat is known (**2**). Special attention is being paid to the transitions from X' to X and from Y to Y' in Figs. 1 and 2, as it will be shown below that these transitions may have occurred during the evolution of the solar system.

11. No attempt will be made in this paper to give a quantitative treatment of the transition of a planet from one configuration to another. Indeed, a detailed treatment is probably impossible at the present time. The scale of the problem, the magnitude of the energy released, and the temperatures and pressures prevailing in the interiors of the planets require dubious extrapolations from the familiar conditions of controlled experiments. A description of the transition is therefore necessarily tentative. Temperature changes need not be considered in a qualitative discussion as they are not of fundamental importance. Indeed, a planet which is initially at the absolute zero of temperature will remain at zero temperature throughout the transition. This is a consequence of the fact that, at the absolute zero, the temperature is not affected by an adiabatic alteration of the pressure. In other cases the temperature will change during the transition, and this will have a small influence which should be taken into account in a quantitative treatment. It has been pointed out above that the central pressures of the initial and final configurations differ by about a million atmospheres. On the other hand, the strength of a solid is only of the order of magnitude of a thousand atmospheres. For these large pressure differences the planet will therefore behave, roughly speaking, as though it were liquid. The time taken for the transition from one configuration to another will be very short. For example, consider the transition from configuration Y to configuration Y', in which the planet collapses and its total radius shrinks by about 100 km. The time taken for the collapse is greater than, but still comparable with, the time taken for a free particle at the surface of the planet to fall through 100 km. under gravity, that is, a matter of minutes or, at

most, hours. In this short space of time an energy of 10^{36} or 10^{37} ergs is liberated. In the transition from Y to Y', in which the planet collapses, this energy is supplied by the gravitational field. In the transition from X' to X, in which the planet expands, the energy is derived from the work done by the internal pressure during the expansion. The energy liberated will, initially, be partly in the form of blast waves and partly in the form of vibrations of the planet as a whole. In a preliminary study of the dynamics of the transitions, Lighthill finds that in the transition from Y to Y', in which the planet collapses, almost the whole of the energy liberated goes into planetary vibrations. But in the transition from X' to X, in which the core disappears, a large fraction of the energy goes into blast waves. The pressure amplitude of these waves will be hundreds of thousands of atmospheres. The blast waves will travel through the planet and be reflected at the outer surface. But the surface of the planet will not be a good reflector of these waves. The material on the surface will be shattered by the sudden impact of such large pressures and fragments may fly off the planet into space. A large fraction of the total energy released may be used to remove fragments from the planet, and so the mass of the material removed may be of the order of magnitude of $0{\cdot}001\ M_E$. It is tentatively suggested that this mechanism may have been responsible for meteorites, and possibly also for the asteroids.

12. This mechanism to account for meteorites and asteroids is based on the assumption that, at some time during the evolution of the solar system, one of the terrestrial planets was in a position to undergo a transition from one configuration to another. It will now be shown that this assumption is not without foundation. So far in this paper little attention has been paid to the thermal conditions prevailing in the interior of the planet. Our knowledge of the thermal histories of the planets, including the Earth, is, at present, very rudimentary. But one point is clear. The temperatures in the deep interior have varied considerably during the planet's evolution, probably by more than 10,000 deg. C. This conclusion is not based on a particular theory of the origin of the solar system. If the planets were formed from material ejected from a star, an enormous decrease in temperature is inevitable. On the other hand, if the planets were originally cold they would be heated by radioactivity. Assuming that the planets have the same overall radioactive content as meteorites, the heat generated during geological time is sufficient to raise the temperature of the planet by more than 20,000 deg. C. Some of this heat will, of course, escape as radiation from the surface. But the major part of the heat will be retained if the planet was originally cold. It may therefore be assumed that the mean internal temperature T of the planet has changed by at least 10,000 deg. C. during its evolution. This change in temperature will have an important bearing on the phenomenon discussed in this paper. The critical masses M_0 and M_1 will be functions of the internal temperature T, and so also of time. The critical mass range will be

$$M_1(T) < M < M_0(T). \tag{38}$$

For any given internal temperature T, the previous treatment is applicable. A planet whose mass M lies in the critical range (38) has three possible equilibrium configurations which are represented by the points A, B and C in Figs. 1 and 2; configurations A and C are stable, and B is unstable. But, as M_0 and M_1 now vary with time, a planet whose mass lies in the critical range at one time need not lie in the critical range at another time. And a planet could cross through

the critical range, either from left to right or from right to left. If a planet crosses the critical mass range in either direction, it will undergo a transition of the type described in the previous paragraph. First, consider a crossing from left to right. Initially the planet's mass is less than the critical mass $M_1(T)$, and its only possible equilibrium configuration lies on the branch YX of Fig. 1. When the planet enters the critical mass range its configuration is represented by the point X. As the internal temperature varies, the planet's configuration effectively moves along the branch XY of Fig. 1, and it reaches Y when its mass M is equal to the critical mass $M_0(T)$. But configuration Y is unstable, and the planet undergoes a transition to configuration Y'. The consequences of this transition are as described in the previous paragraph. Second, suppose that the planet crosses the critical range (38) from right to left. Initially its mass is greater than the critical mass $M_0(T)$, so that its configuration lies on the branch $X'Y'$ of Fig. 1. As the internal temperature varies, the planet's configuration effectively moves towards the point X', and it reaches X' when its mass M is equal to the critical mass $M_1(T)$. But configuration X' is unstable, and the planet undergoes a transition to configuration X. Thus both of the planetary transitions described in the previous paragraph could be realized in nature.

A reasonable variation in the internal temperature, such as 10,000 deg. C., can change the critical masses M_0 and M_1 only by a small percentage. But the change is nevertheless larger than the mass difference $(M_0 - M_1)$. For example, consider the critical mass M_0. According to equation (2), M_0 can be written in the form:

$$M_0 = (6/\pi)^{1/2}(p_c/G)^{3/2}\rho_0^{-2}, \tag{39}$$

where the critical pressure p_c and the density ρ_0 are functions of the internal temperature T. Unfortunately, the relationship between critical pressure and temperature is unknown. This situation is inevitable when nothing is known about the latent heat associated with the phase transition at the boundary of the core; the Clausius-Clapeyron equation connects this latent heat with the temperature derivative of the critical pressure. One is not even in a position to say if the critical pressure is an increasing or a decreasing function of the temperature. Kronig, de Boer and Korringa (**7**) estimate that a change in temperature from absolute zero to 5000 deg. C. would not have a noticeable effect on the critical pressure for the formation of metallic hydrogen. This is probably true of every transition to a metallic phase at very high pressures. But a change in the critical pressure of a few per cent on increasing the temperature by 10,000 deg. C. would not be surprising. Temperature probably influences the critical mass M_0 mainly through the density ρ_0, which is altered by thermal expansion. The coefficient of cubic expansion of common rocks under laboratory conditions is about $2{\cdot}5 \times 10^{-5}$ per deg. C. The coefficient of thermal expansion increases slightly with temperature, but it decreases rapidly with pressure. At the temperature and pressure at the boundary of the Earth's core the coefficient of thermal expansion is probably an order of magnitude smaller than under laboratory conditions. Nevertheless, a change in temperature of 10,000 deg. C. could alter the effective value of the density ρ_0 by 5 per cent, and so the critical mass M_0 by 10 per cent. For the present purposes it is not essential to know precisely by how much the critical masses M_0 and M_1 have changed during geological time. It is sufficient to know that the changes have been appreciable.

13. For the terrestrial planets the critical mass M_0 is about 0·80 M_E, M_E being the mass of the Earth (*cf.* equation (20)). This is smaller than the mass of Venus, which is 0·817 ± 0·005 M_E. But the difference is only 2 per cent, and M_0 could easily have changed by this amount in the course of geological time. Venus might therefore have undergone a transition of the type envisaged. If the mass of Venus has crossed the critical range (38), it must have been in the direction from left to right, as the mass of the planet is now greater than the critical mass M_0. This would lead to a transition of the type YY' in Figs. 1 and 2, in which the planet collapses and a core is formed. Lighthill estimates that almost the whole of the energy released by this type of transition goes into radial vibrations of the planet. This energy will not be available to remove material from the surface of the planet. If the whole of the energy released were utilized to remove fragments from the surface of the planet, the mass of material removed would be of the order of magnitude of 0·001 M_E. But the existing evidence suggests (**8**) that the aggregate mass of the asteroids is also about 0·001 M_E. It therefore seems certain that the planetary transition suggested for Venus could not account for the asteroids. But this transition might account for meteorites. It is an open question whether all meteorites belong to the solar system, but let us suppose that they do. The total annual amount of meteoric material collected by the Earth is estimated to be about 360 tons (**8**). Assuming that meteorites are uniformly distributed over a sphere of the same dimensions as the solar system, it follows that the aggregate mass of meteorites is only about one-millionth of the mass of the Earth. During geological time the planets have swept up only a small fraction of the total number of meteorites. The planetary transition suggested for Venus could account for meteorites, even if only 0·1 per cent of the energy released was available to remove fragments from the surface of the planet.

A planetary transition of the type $X'X$, in which the core disappears, would be required to account for the asteroids. The possibility of such a transition cannot be excluded. In addition to the main core, the Earth also possesses an inner core about 1200 km. in radius. The author (**2**) has explained the inner core in terms of a second phase transition of the type responsible for the core. Little is known about the jump in density at the boundary of the inner core; it may or may not exceed 50 per cent. If it does exceed 50 per cent, a planet could undergo a transition of the type $X'X$, in which the inner core disappears. Because of their masses, Venus and Pluto are the most likely candidates for this particular planetary transition. There is also the possibility that the Earth has lost a third innermost core.

Note added in proof.—Kuiper (**9**) has recently resolved the disk of Pluto for the first time, and he estimates its radius to be about one-half of that of the Earth. This radius corresponds to a mass of 0·07 M_E, which is an order of magnitude less than the previous estimates based on the perturbations of Neptune. The central pressure in a planet of this mass could not support a central core, and a planetary transition of the type envisaged is therefore out of the question.

14. The idea of an " exploded " planet or a planet " spoiled in the making " has been current for about a century. This has been considered necessary to account for meteorites. It is also suggested by the irregular shapes and rough surfaces of the asteroids. But this view could not be accepted, as there was no known reason why a planet should explode. The mechanism suggested in the present paper overcomes this difficulty. The mechanism is both possible and

plausible. To this extent the old idea of an " exploded " planet has been given substance. But one cannot assert that either asteroids or meteorites did, in fact, originate in this way. The extreme complexity of the problem prevents a complete quantitative treatment which might give a definite answer. The present work should be regarded as re-opening, rather than closing, the question of the origins of meteorites and asteroids.

The Physical Laboratories,
The University,
Manchester, 13:
1950 *April* 11.

References

(1) W. H. Ramsey, *M.N.*, **108,** 406, 1948.
(2) W. H. Ramsey, *M.N., Geophys. Suppl.*, **5,** 409, 1949.
(3) K. E. Bullen, *An Introduction to the Theory of Seismology*, Cambridge, 1947.
(4) Harold Jeffreys, *The Observatory*, **70,** 22, 1950.
(5) H. Poincaré, *Figures d'équilibre d'une masse fluide*, Paris, 1903.
(6) J. H. Jeans, *Astronomy and Cosmogony*, Cambridge, 1928, ch. 8.
(7) R. Kronig, J. de Boer and J. Korringa, *Physica*, **12,** 245, 1946.
(8) H. N. Russell, R. S. Dugan and J. Q. Stewart, *Astronomy*, Boston, 1945.
(9) G. P. Kuiper, *P.A.S.P.*, **62,** 133, 1950.

ON THE INSTABILITY OF SMALL PLANETARY CORES (II)

M. J. Lighthill

(Communicated by W. H. Ramsey)

(Received 1950 April 13)

Summary

The property that, for a spherically symmetrical planet in which the density is a function of the pressure, three states of equilibrium are possible in a certain range of values of the total mass, is shown to hold whenever the density is continuous up to a critical pressure p_c, at which (owing to a change of phase) it rises discontinuously by a factor exceeding $\frac{3}{2}$. The question of transitions between the states is briefly discussed.

---

1. *Introduction.*—For a spherically symmetrical planet at rest, in which the density ρ is a specific function $\rho(p)$ of the pressure p, the equation

$$\frac{dp}{dr} = -\rho(p)\frac{G}{r^2}\int_0^r \rho(p) 4\pi r^2 dr \tag{1}$$

(in which G is the gravitational constant and r represents distance from the centre) determines the pressure distribution $p(r)$ uniquely in terms of the pressure at the centre $p(0)=p_0$. Hence the whole configuration is determined when p_0 is known, including the mass M of the planet given by the equation

$$M = \int_0^R \rho(p) 4\pi r^2 dr = -\frac{R^2}{G\rho(0)}\left(\frac{dp}{dr}\right)_{r=R}, \tag{2}$$

in which R is the least value of r for which $p=0$ (i. e. the radius of the planet).

But it was shown in Part I that for a certain equation of state ($\rho=\rho_0$ for $0<p<p_c$, $\rho=\lambda\rho_0$ for $p>p_c$, with $\lambda>\frac{3}{2}$) the mass M does not determine the pressure at the centre p_0 uniquely, but that for some values of M three possible configurations of the planet exist, of which either the one with largest p_0 or the one with smallest p_0 is the most stable. Planets with very small cores are in the least stable of three such configurations.

It will here be shown that for any equation of state which is continuous up to a critical pressure p_c, where a phase change causes a jump in density by a factor exceeding $\frac{3}{2}$, there will be three statically possible configurations for certain values of the mass M.

The precise result to be proved is as follows. *If $\rho(p)$ is continuous for $0\leqslant p\leqslant p_c$ but takes a value $\lambda\rho_1$ (where $\rho_1=\rho(p_c)$ for p just greater than p_c, then in the dependence of the mass M on the central pressure p_0 the gradient dM/dp_0 is changed discontinuously by a factor $(3-2\lambda)/\lambda^2$ at $p_0=p_c$.* The same will be true of the gradient, with respect to p_0, of any other physical quantity (such as the radius R) which depends on the bulk properties of the planet. For it will be shown that, for $p_0=p_c+\delta$, where δ is small, the configuration, at distances from the centre large compared with the radius of the (small) core, is the same as if no change of phase had occurred (with the continuous equation of state $\rho=\rho(p)$ extrapolated if necessary beyond $p=p_c$) and as if the central pressure were $p_c+[(3-2\lambda)/\lambda^2]\delta$.

From the italicized result it follows that for $\lambda > \frac{3}{2}$ the gradient dM/dp_0 changes sign. Hence, if, on the one hand (as is presumably normal), M increases with p_0 for p_0 just less than p_c, then M must decrease as p_0 rises above p_c. Assuming that finally M must increase again, i.e. $M \to \infty$ as $p_0 \to \infty$ (because it is physically unacceptable that a finite mass could statically support an infinite pressure), there must be a band of values of M, bounded above by the value M_0 given by $p_0 = p_c$, for which at least three configurations exist—one with no core (i.e no material with $p > p_c$), and two with cores, one smaller and one larger.

Poincaré's familiar arguments, which he incorporated in his "principle of exchange of stabilities", make it clear that the inner of these three states of equilibrium (the one with the smaller core) is unstable and that the outer two are stable, because the only possible form of the energy curves is as indicated in Fig. 2 of Part I. As the critical mass range $M_1 < M < M_0$ alters by change of temperature, the central pressure could approach the critical value p_c (so that M approached M_0) from below through stable states of equilibrium, but after this the planet would make a sudden transition to the only state of equilibrium with a mass just exceeding M_0. This state has already a non-small core. Conversely a planet with a core might slowly alter its state until its mass reached the value M_1 at which the core radius has shrunk to the instability point, and then in a sudden transition the core would disappear, and the stable state without a core would appear. The dynamics of these transitions is being investigated. Preliminary considerations indicate that only in the transition where a core disappears does the energy released become (at any stage) so concentrated locally as to make it likely that a substantial part of it could go into release of matter from the planet's gravitational field.

If, on the other hand, M is already decreasing (as a function of p_0) for p_0 just less than p_c (it is possible that this might occur—for example, if the equation of state up to $p = p_c$ included a rapid but not discontinuous jump by a factor exceeding $\frac{3}{2}$), then it must immediately increase again as p_0 rises above p_c. Since M originally increases from zero at $p_0 = 0$, there must in this case be a band of values of M, bounded *below* by M_0, for each of which at least three configurations exist—two with no core and one with a core. In this case a small core would presumably be stable.

2. *Proof of the results.*—The solution of equation (1) with $p_0 = p_c$ is assumed known, say $p = P(r)$. An approximate solution is sought with $p_0 = p_c + \delta$, where δ is small.

Then in the small core in which $p > p_c$ (and therefore $\rho \simeq \lambda\rho_1$) equation (1) becomes approximately

$$\frac{dp}{dr} = -\lambda\rho_1 \frac{G}{r^2} \lambda\rho_1 \tfrac{4}{3}\pi r^3 = -\tfrac{4}{3}\pi G\lambda^2\rho_1^2 r. \tag{3}$$

Hence by integration $p = p_c + \delta - \frac{2}{3}\pi G\lambda^2\rho_1^2 r^2$, and the radius of the core r_c is given by the condition $p(r_c) = p_c$ as

$$r_c \simeq \frac{1}{\lambda\rho_1}\sqrt{\left(\frac{3\delta}{2\pi G}\right)}. \tag{4}$$

The conditions for r just greater than r_c may now be written down (remembering that here $\rho = \rho_1$, so that (3) is not quite the right approximation to (1)) as

$$p = p_c, \quad \frac{dp}{dr} = -\rho_1 \frac{G}{r_c^2} \lambda\rho_1 \tfrac{4}{3}\pi r_c^3 = -\tfrac{4}{3}\pi G\lambda\rho_1^2 r_c. \tag{5}$$

For $r>r_c$ equation (1) will be used in the differentiated form

$$\frac{d}{dr}\left(\frac{r^2}{\rho(p)}\frac{dp}{dr}\right)=-4\pi Gr^2\rho(p), \tag{6}$$

and in this region $\rho(p)$ is a continuous function. The boundary conditions (5) holding as $r\to r_c$, determine a unique solution to equation (6) in $r>r_c$.

Now since these boundary conditions differ only slightly from those ($p=p_c$ and $dp/dr=0$ at $r=0$) which determine the solution $P(r)$, it is permissible to put $p=P(r)+\Pi(r)$ in (6) and deduce a linear equation for $\Pi(r)$ by neglecting squares of $\Pi(r)$ (or products of itself with its derivatives). Thus we may write

$$\frac{d}{dr}\left[r^2\frac{P'(r)+\Pi'(r)}{\rho(P(r))+\Pi(r)\rho'(P(r))}\right]=-4\pi Gr^2[\rho(P(r))+\Pi(r)\rho'(P(r))]\,; \tag{7}$$

and, remembering that $P(r)$ itself satisfies (6), and collecting the terms in (7) which are linear in Π (while discarding quadratic ones), deduce that

$$\frac{d}{dr}\left[r^2\left\{\frac{1}{\rho(P)}\frac{d\Pi}{dr}-\frac{P'\rho'(P)}{\rho^2(P)}\Pi\right\}\right]=-4\pi Gr^2\rho'(P)\Pi. \tag{8}$$

The only fact needed about this linear equation for Π is that it possesses two linearly independent solutions, of the form

$$\left.\begin{aligned}\Pi^{(1)}&=r^{-1}+\Pi_1^{(1)}r+\Pi_2^{(1)}r^2+\ldots,\\ \Pi^{(2)}&=1+\Pi_1^{(2)}r+\Pi_2^{(2)}r^2+\ldots.\end{aligned}\right\} \tag{9}$$

To prove this it may be verified that the equation for $y=r\Pi/\rho(P)$ takes the considerably simpler form

$$\frac{d^2y}{dr^2}+4\pi G\rho(P)\rho'(P)y=0, \tag{10}$$

in which the coefficient of y is finite for all r since $P(r)\leqslant p_c$ for all r. Hence y possesses a solution with arbitrary values of $y(0)$ and $y'(0)$, which general solution to the linear equation (10) may be put in the form

$$r\Pi=\rho(P)y=A[1+\Pi_1^{(1)}r^2+\Pi_2^{(1)}r^3+\ldots]+B[r+\Pi_1^{(2)}r^2+\Pi_2^{(2)}r^3+\ldots], \tag{11}$$

giving linearly independent solutions for Π as in (9).

The combination of solutions (9) which is here required is determined by the boundary conditions. Since at $r=r_c$

$$\left.\begin{aligned}P+\Pi&=p_c,\quad P'+\Pi'=-\tfrac{4}{3}\pi G\lambda\rho_1^2r_c,\\ P&=p_c-\tfrac{2}{3}\pi G\rho_1^2r_c^2,\quad P'=-\tfrac{4}{3}\pi G\rho_1^2r_c\end{aligned}\right\} \tag{12}$$

(by (5), and by approximating equation (1) for P under the condition $P(0)=p_c$), the boundary conditions on Π are that, at $r=r_c$,

$$\Pi=\tfrac{2}{3}\pi G\rho_1^2r_c^2,\quad \Pi'=\tfrac{4}{3}\pi G(1-\lambda)\rho_1^2r_c. \tag{13}$$

Hence if $\Pi=A\Pi^{(1)}+B\Pi^{(2)}$, so that $\Pi\simeq Ar^{-1}+B$ for small r, we have approximately

$$\frac{A}{r_c}+B=\tfrac{2}{3}\pi G\rho_1^2r_c^2,\quad -\frac{A}{r_c^2}=\tfrac{4}{3}\pi G(1-\lambda)\rho_1^2r_c, \tag{14}$$

whence, using (4),

$$A = \tfrac{4}{3}\pi G(\lambda - 1)\frac{1}{\lambda^3 \rho_1}\left(\frac{3\delta}{2\pi G}\right)^{3/2}, \tag{15}$$

and

$$B = \tfrac{2}{3}\pi G\rho_1^2 r_c^2 + \tfrac{4}{5}\pi G(1-\lambda)\rho_1^2 r_c^2$$
$$= \frac{3-2\lambda}{\lambda^2}\delta. \tag{16}$$

Now, for small δ, A is small compared with B. On the other hand, $\Pi^{(1)}$ and $\Pi^{(2)}$ are independent of δ and are ordinary finite functions when r is not small. Hence the configuration, outside a region $r = O(\delta^{1/2})$ *, is given approximately by $\Pi = B\Pi^{(2)}$, i. e.

$$p = P(r) + B\Pi^{(2)}(r). \tag{17}$$

Now *this* is a solution not only of the differential equation (6) but also of its integrated form (1), since $r^2 dp/dr$ is zero at $r = 0$. Hence it is simply a solution of the problem of a planet with $p_0 = p_c + B$ and with the continuous equation of state $\rho = \rho(p)(0 \leqslant p \leqslant p_c)$ extrapolated beyond $p = p_c$ if necessary.

In other words, as stated in the introduction, the configuration at distances from the origin large compared with the radius of the core is the same as if no change of phase had occurred and the central pressure were $p_c + [(3-2\lambda)/\lambda^2]\delta$.

In particular the radius R, where $p = 0$, and the mass M given by (2), will be the same as if this were so. Thus

$$M \simeq M_0 + \frac{3-2\lambda}{\lambda^2}\delta\left[\frac{dM}{dp_0}\right]_{p_0 = p_c - 0}. \tag{18}$$

Hence dM/dp_0 changes discontinuously at $p_0 = p_c$ by a factor $(3-2\lambda)/\lambda^2$.

As shown in the introduction this implies that three configurations are statically possible, in a certain range of values of the total mass, whenever $\lambda > \frac{3}{2}$.

Department of Mathematics,
The University, Manchester, 13:
1950 *April* 11.

* Wherein the r^{-1} in $\Pi^{(1)}$ makes $A\Pi^{(1)}$ comparable with $B\Pi^{(2)}$.

OBSERVATIONS OF THE BRIGHTEST REGIONS OF THREE DIFFUSE NEBULAE

A. D. Thackeray

(Received 1949 September 26)

Summary

Observations, by direct photography and spectroscopy, of the three diffuse nebulae M16, M20 and M8 are presented and discussed, with special reference to the bright rims around associated dark lanes.

The spectra of the rims are found to consist of a strengthening of the nebular spectrum. The ratio $H\beta$/N1 is greater at or near the rims. Weak continuous spectrum is found in the brightest rim of M16.

Qualitative evidence is presented that the dark matter is moving inwards towards a central star or stars against the bright nebulosity.

The brightest part of M8 reveals a complex structure intimately associated with faint stars. A green photograph, avoiding emission lines, reveals nuclei which may be stellar. There is a strong continuous spectrum, in addition to a nebular spectrum with permitted lines enhanced relative to forbidden.

---

1. This paper describes observations with the 74-inch Radcliffe reflector of the three well-known diffuse nebulae M16 (NGC 6611), M8 (6523) and of the Trifid M20 (6514). The main purpose of the investigation was to study the bright rims which are observed surrounding the dark lanes which are projected against these diffuse nebulae. M16 is probably the best example of a nebula rich in such rims, of which the "Horse's Head" (IC 434) is a familiar instance. The Eta Carinae nebula also displays a number of good examples, but it is such an extensive and complex object that further observations are required before they can be discussed.

The observations have been made (*a*) by direct photography at the Newtonian focus (scale 22·5 seconds of arc per mm.) with various emulsions and colour filters, and (*b*) spectroscopically, also at the Newtonian focus. The spectrograph employed is a fast two-prism instrument with collimating mirror and $f/1$ Schmidt camera by Cox, Hargreaves and Thomson Ltd.; the spectrograph box was made in the workshops of the Observatories, Cambridge, with the kind cooperation of Professor Redman, who also assembled and tested the spectrograph at Cambridge before it was despatched to Pretoria in 1949 April. The writer's thanks are due to Messrs Cox and Hargreaves and to Professor Redman for the speed with which work on this instrument was completed, which enabled it to be in operation about a year after agreement on the final specifications. The dispersion of the spectrograph is 240 A./mm. at $H\gamma$ and spectra of fine definition can be obtained on film disks, of about 8 mm. diameter, covering the range 5900 to 3727 A. The original film-retaining ring was modified to extend the region so as to include the important nebular lines at 3727 A. which appear frequently despite the rather heavy absorption at this wave-length in the glass prisms. No attempts at radial velocity measurements in these diffuse nebulae have been made, and in fact no comparison spectrum was normally impressed, since it was rather important that the whole length of the spectrum lines should be available for study.

In Table I are listed details of the various direct photographs and spectra of the three nebulae. In the following sections the results of measures of selected portions of the three nebulae are presented separately.

TABLE I

Observations of M16, M20 and M8

Direct Photographs:

Plate	Date (1949)		Emulsion	Filter	Exposure (mins.)	Remarks
A 371	July 21	M16	I.A.Z.	...	30	
372	21	M16	103 AE	W 12	30	Plate 4 (*a*).
383	28	M16	103 AO	...	20	Calibrated.
384	28	M16	103 AE	W 15	40	Calibrated.
391	Aug. 13	M8	103 AE	W 12	30	Centre blacked out.
395	20	M16	103 AO	W 18A	40	Ultra-violet.
402	25	M20	O–E	...	10	Plate 5 (*a*).
411	27	M20	103 AO	W 18A	30	Ultra-violet.
412	27	M20	O–F	W 58	30	Green; no nebula shown.
413	29	M8	O–E	W 25	5	Red; Plate 6 (*a*).
414	29	M8	O–F	W 58	30	Green; Plate 6 (*b*).
415	29	M16	103 AO	W 58	25	Green; no nebula shown.
422	31	M8	103 AO	W 58	15	Nebula (N + continuous) shown.
423	31	M8	O–E	W 25	1½	Red; nebula shown.

Spectra (Emulsion: R55, orthochromatic):

Film	Date (1949)	Slit (mm.)	Exposure (mins.)	Object	Remarks
NS 72	July 30	0·15	40	M16; rim A	
73	30	0·15	60	M16; rim B	
76	Aug. 1	0·15	120	M16; rim C	
79	3	0·10	100	M16; rim B	
94	15	0·12	90	M16; rim A	Plate 4 (*b*).
95	15	0·15	60	M16; region D	
101	22	0·75	60	M16; rim A	Wide slit; continuous spectrum.
102	22	0·15	30	M8	Across Herschel 36.
107	25	0·15	30	M20	Slit EW; brightest part.
108	27	0·15	50	M20	Slit p.a. 72° crossing rim; Plate 5 (*b*).
109	29	0·15	60	M8	10″ S. of Herschel 36; Plate 6 (*c*).

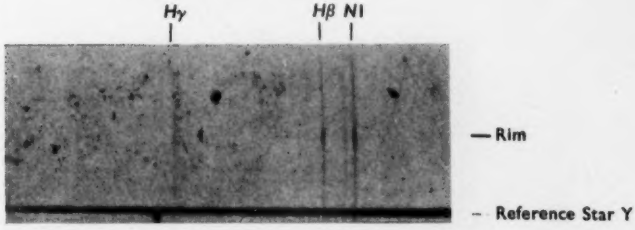
M16 (NGC 6611) (Plate 4 and Fig. 1)

This remarkable nebula and cluster have been described by Duncan.* Perhaps its most outstanding feature is the central dark area divided into three roughly parallel portions which point in the S.E.–N.W. direction towards the cluster. Each of these dark lanes ends abruptly in a bright rim whose brightest portion lies nearest to the cluster. The most northerly appears to be partly veiled by bright foreground nebulosity near the star which lies in a bay between this and the next dark lane. This star, which we denote X, and which serves as a reference point in these observations of M16, lies 51″ preceding and 72″ south of the star AG 6241.† The brightest area in the whole nebula, lying about 1′ of arc N.W. of X, does seem to mark the end of the northern lane, and this area which we denote "region A" has received most attention. Region B refers to the bright tip of the

* J. C. Duncan, *Ap. J.*, **51**, 4, 1920; and *Mt. Wilson Contr.*, No. 177, 1920.

† *Trans. Yale Obs.*, **14**, 1943. SDM −13° 4932 in Duncan's list of stars.

(*a*) *M*16. *Scale* 6″·9 = 1 *mm.*



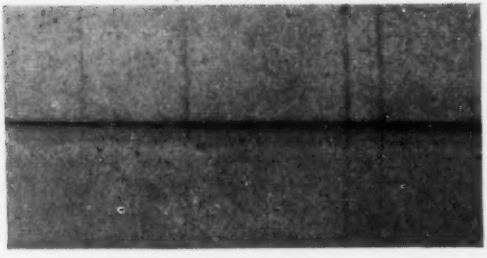
(*b*) *Rim A and reference star. Scale along lines same as for* (*a*), *east at top.*

A. D. Thackeray, *Observations of the Brightest Regions of Three Diffuse Nebulae.*

S

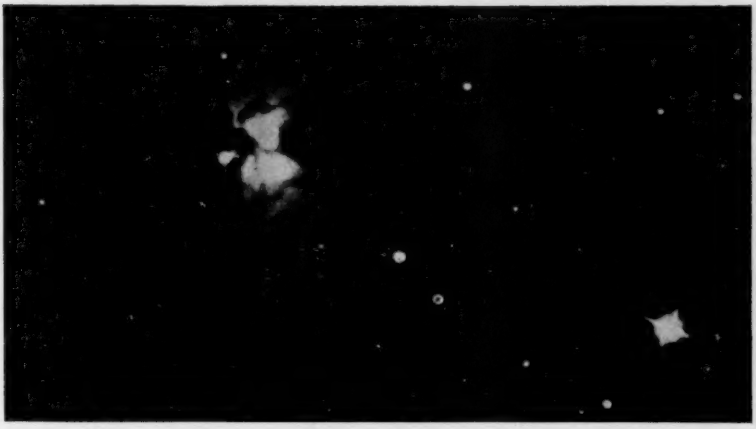
E

(*a*) *M*20. *Scale* 2″·7 = 1 *mm.*



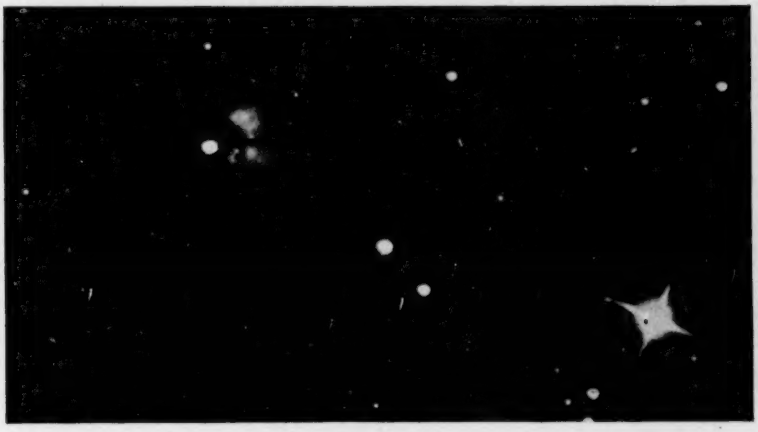
(*b*) *Spectrum of rim with slit orientated p.a.* 72° *as indicated in Fig.* 4.
Scale along lines : 6″·35 = 1 *mm., east at top.*

S

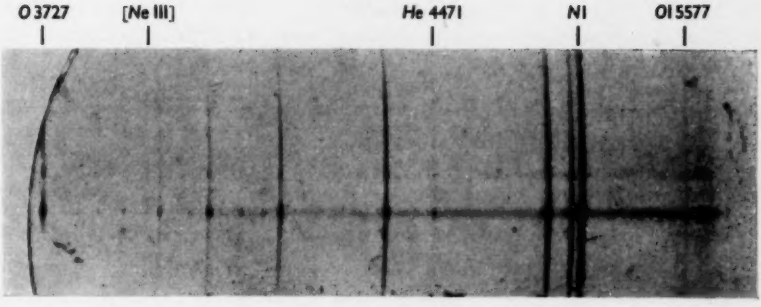


E

(*a*) *Red. M8. Scale* 2″·1 = 1 *mm.*



(*b*) *Green.*



(*c*) *M8. Spectrum of southern half of nucleus.*
Scale along lines : 6″·3 = 1 *mm., west at top.*

A. D. Thackeray, Observations of the Brightest Regions of Three Diffuse Nebulae.

second dark lane almost due west of X; region C refers to the bright rim on the northern side of the dark circular patch 6′ south of the cluster. (Fig. 1 shows the positions of the slit as used in these various regions.)

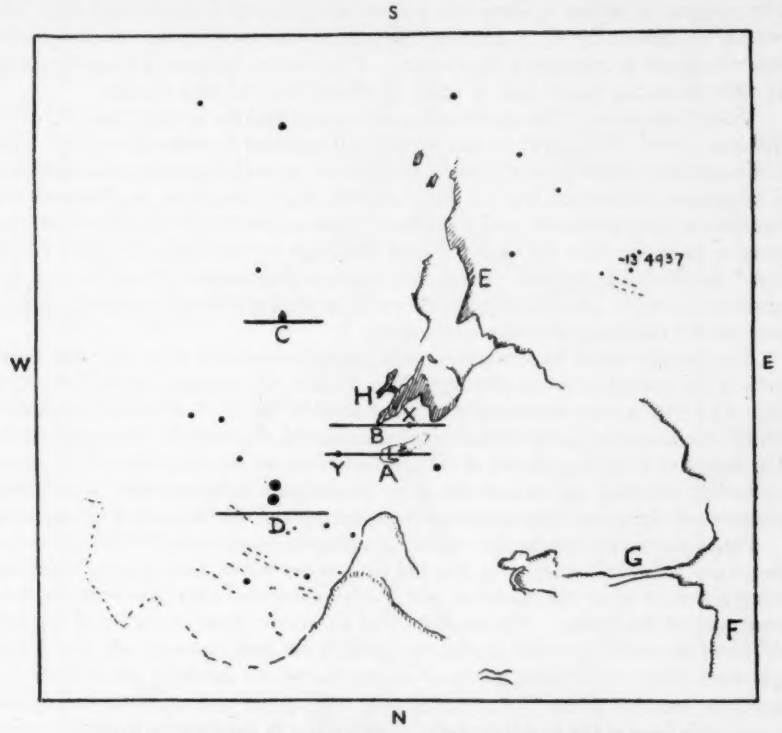


FIG. 1.—*Sketch of M16 showing positions of slit in regions A, B, C and D. Transitions from bright to dark areas are outlined; shadings indicate the dark side to emission rims.*

The approximate coordinates of the centre of the slit relative to the star X where spectra of these various regions were obtained are given in Table II. The slit was always set due east–west partly because this facilitated more precise definition of where the slit actually lay, and partly because more precise guiding could be attained. There is at present no means of viewing the slit during exposures, and guiding was done by means of one of the offset eyepieces normally

TABLE II

Coordinates of regions in M16 studied spectroscopically relative to star X and positions in millimetres from edges of Plate 4

Region	$\Delta\alpha$	$\Delta\delta$	Position in Plate 4	
X	0	0	R50	B56
A	− 42″	+ 45″	R56	B50
Ref. Star Y near A	−126	+ 45	L46	B50
B	− 60	− 6	L56	B57
C	−221	−189	L32	T52
D	−240	+200	L30	B35

used on the Newtonian breechpiece. But with careful setting the position of the slit can be defined to within a few seconds of arc, especially if there is a star as bright as 12^m at hand which can be viewed directly through the spectrograph. The spectra of region A show the narrow spectrum of a moderately faint star (which we denote Y) 84″ preceding the bright rim and this served as a useful reference point in measuring the spectra. The spectra of region B could similarly be referred to the bright star X, some of whose light entered the slit.

Visual inspection of the spectra showed at once that the nebular lines N1, N2, $H\beta$ and γ, and [O II]3727 (when shown) all increase in intensity at the rims. Although these small films are hardly suitable for microphotometry, and no means of laboratory calibration has yet been devised, they have been run through the Radcliffe microphotometer and an attempt made to study the intensity variations, using as basis the ratio $N1/N2 = 3{\cdot}7$, this being the average value found by Baade *et al.** for the Orion nebula. Runs were made at a succession of points along the spectrum lines N1, N2, $H\beta$, $H\gamma$, and the variation of peak intensity at corresponding parts of the nebula could thus be studied.

The results could be compared with microphotometer runs over the same parts of the nebula on direct photographs. Plate A 383, a rather dense blue plate, calibrated with a tube sensitometer (constructed by Dr D. S. Evans, who has also kindly communicated to me the calibration constants), was used for this comparison. The length of the analysing slit of the microphotometer corresponded to 0·15 mm. on the film or plate; the runs on the direct photograph therefore refer to the same width of nebula as was transmitted by the spectrograph slit for most of the spectra.

Fig. 2 shows the results for region A in diagrammatic form. The top curve shows the intensity variation across the rim on the direct photograph, abscissae being distances from the reference star Y whose position coincides with the left-hand edge of the figure. The second curve shows the mean variation of N1 and $H\beta$ from the two films (NS 94 and 101) which are best exposed and show long spectrum lines. The third group of curves shows the intensity ratio $H\beta$/N for the three separate films, NS 72 being only sufficiently well exposed to provide measurable lines at the brightest part; in this ratio, N denotes the weighted mean $(3N1 + N2)/4$, the measures of N2 being frequently rather unreliable. Finally, in the lowest curve of Fig. 2 is shown the average photometer deflection in millimetres from five runs, ***transverse to the spectrographic dispersion***, at various wave-lengths near $H\beta$ but avoiding emission lines. This curve is derived from NS 101 only; this spectrum was obtained with a very wide slit (0·75 mm.) to test the presence of continuous spectrum which had already been suspected on NS 94.

Fig. 3 shows curves similar to those of Fig. 2 (omitting the last) derived from region B.

From Fig. 2 we may draw the following conclusions regarding the bright rim A:

(1) The nebular lines N1, N2 and $H\beta$ all increase to a sharp maximum of intensity at the rim.

(2) The intensity at the rim is about 2·5 times that of the surrounding nebulosity both in the nebular lines and in integrated blue light.

(3) There are indications of an abrupt change in the intensity-ratio $H\beta$/N at or near the rim in the sense that $H\beta$ is relatively strongest at the rim. (In comparing the separate results of the three films it must be remembered that NS 101, with

* W. Baade, F. Goos, P. P. Koch and R. Minkowski, *Zs. f. Ap.*, **6**, 355, 1933.

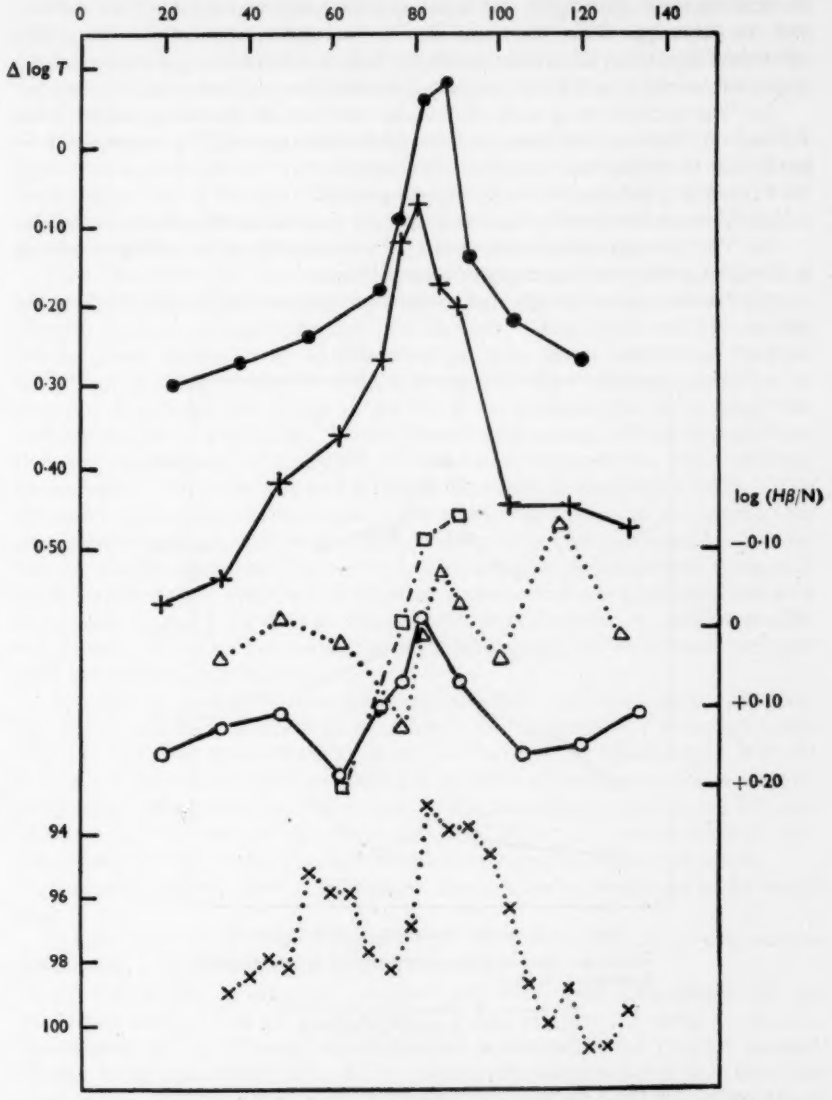


FIG. 2.—*Variation of intensity in M16 (region A).*

Abscissae : distance from star Y in seconds of arc.
Ordinates (lowest graph): linear deflection in millimetres.

Δ *log I* ●———● *Direct photograph (blue).*
Δ *log I* +———+ *Mean of N1 and Hβ.*
log (*Hβ*/*N*) ○———○ *NS* 101 (*wide slit*).
log (*Hβ*/*N*) Δ———Δ *NS* 94.
log (*Hβ*/*N*) □—·—·—□ *NS* 72.
Continuous background ×·····× *Linear deflection of microphotometer.*

slit five times as wide as for the other spectra, embraces more of the nebula; and the other two films may not refer to identical regions of the rim. The difference in position of maximum intensity of N and $H\beta$ was apparent on visual inspection of the film NS 72, and may be detectable on Plate 4(*b*).)

(4) The presence of a weak continuous spectrum at the rim is evident from NS 101. (The secondary maximum in the bottom curve of Fig. 2 appears to be partly due to defects and partly to a faint star.)

From Fig. 3 we may conclude, regarding rim B:

(5) The nebular lines N and $H\beta$ all increase to a maximum intensity at the rim.

(6) The intensity at the rim is about 1·6 times that of the surrounding nebulosity in the nebular lines and in integrated blue light.

(7) There is some indication of a very slight increase in the ratio $H\beta$/N at the rim.

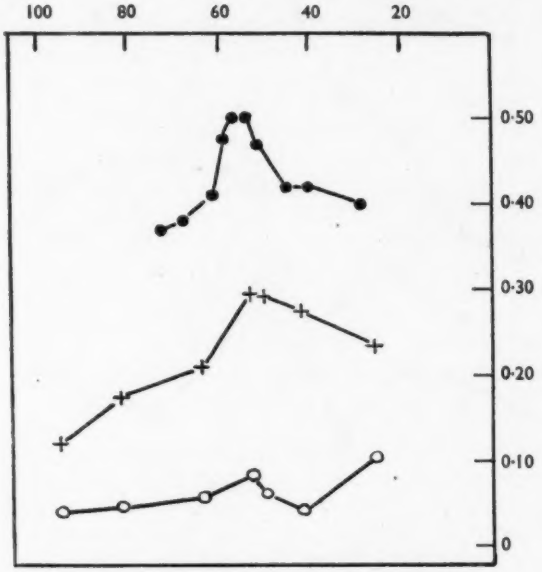


FIG. 3.—*Variation of intensity in M16 (region B).*

Abscissae: distance from reference star X in seconds of arc.
Ordinates: Δ log I.

●———● *Direct photograph (blue).*
+———+ *Mean of N1 and Hβ.*
○———○ *log (Hβ/N).*

In examining these curves various points must be borne in mind.

(*a*) The reduced scale in the spectrograph along the lines is 95″ per mm.; he width of the bright rims amounts to only a few seconds of arc, and even with the slit set tangential to the rim as for B, the analysing slit of the photometer is not short enough to show as sharp a peak in the spectra as in the direct photographs. Thus the measured intensity ranges in the spectra are probably underestimated. On the other hand, the measured intensities in the spectra some way from the maxima are subject to an unknown degree of vignetting in the spectrograph, and this will tend to increase the apparent range of intensity.

(*b*) In region A the slit ran roughly radial to the rim, which is broader than in the typical case, and the right-hand side of the curves refers apparently to foreground nebulosity veiling a dark lane. In region B the slit was tangential to the rim and both right- and left-hand sides of the curves refer to nebulosity on the bright side of the rim and dark lane.

The single two-hour spectrum of the rim in region C was not very successful, apparently because the slit was set slightly inside the dark patch instead of crossing the brightest part of the rim. Nevertheless, the nebular lines N1, N2 and $H\beta$ are brightest on each side of the dark patch and thus afford similar evidence that this rim represents an enhancement of the typical nebular spectrum.

The conclusion that continuous spectrum is present at the rim in region A is important in that it raises the question whether line or continuous spectrum is more effective in the direct photographs. As Hubble* has pointed out, line spectra are far more readily detected in diffuse nebulae than continuous spectra. Unfortunately, despite the use of such a wide slit, the continuous spectrum in region A is so weak that it lies on the toe of the characteristic curve where the calibration is most unreliable. Nevertheless it may be said with some confidence that the continuous spectrum on NS 101 amounts to between 0·04 and 0·02 times the intensity of N1 at the rim, and probably lies nearer to the smaller value. The slit width in this case corresponded to 72 angstroms where the continuous was measured, or about 0·06 × (total spectrum effective on blue photographs). If we assume that N1 contributes one-quarter of the total line emission effective on blue photographs, we find that the total effective continuous at rim A probably amounts to between 16 and 8 per cent of the total effective line emission, with the smaller value preferred. In other words, the direct photographs of rims almost certainly refer primarily to line emission.

Attempts to isolate the rims monochromatically on direct photographs have not shown any clear changes in structure. An ultra-violet photograph with Wratten 18A filter shows the rims as well as blue and red photographs, but the continuous spectrum at the head of the Balmer series probably contributes as much to the ultra-violet photograph as *O* 3727. The combination Kodak 103 AO plus Wratten filter 58 cuts out $H\beta$ relative to N1 and N2 by a factor of at least 5, but the attempt to record the nebula thus failed owing to inadequate exposure.

The only marked differences between blue and red photographs so far noted are:

(1) the presence of a large number of faint red stars in the cluster—they may be intrinsically red or reddened by interstellar matter;

(2) the relative prominence on blue and ultra-violet photographs of the nebulosity surrounding the cluster, which tends to show the streaky structure reminiscent of the Pleiades nebulae. The direction of the roughly parallel streaks is in p.a. about 60°. It is a remarkable circumstance that the blue photographs also show signs of a similar streaky structure near the 10th mag. star SDM −13° 4937†, which is bluer than most stars in the field, and that though this star lies on the outer side of the main nebula the orientation of the streaks is roughly in the same p.a.

Since it was suspected that this streaky structure arises mainly from the pure reflection type of nebulosity, one spectrum was taken in " region D ", 50″ of arc

* E. Hubble, *Ap. J.*, **56**, 162, 1922.

† See list of stars accompanying Duncan's photograph, *loc. cit.* and Fig. 1. This region lies outside Plate 4 (*a*). The streaks are indicated roughly by dashed lines in Fig. 1.

north of the bright star SDM $-13^{\circ}4925$ (HD 168076) near the centre of the cluster. Here the nebulosity on a blue photograph, as measured with the microphotometer, has about the same intensity as that near (but not on) rim A. The nebular lines at D appear in a 60-minute exposure rather weaker than those near rim A in a 40-minute exposure. Hence, although the slit was too narrow to record a continuous spectrum at D, it appears probable that the light here also consists of a mixture of line plus continuous emission with the continuous providing a considerably greater contribution than at A.

The Trifid Nebula M20 (NGC 6514) (Plate 5 and Fig. 4)

The present observations of this nebula are restricted to the central brightest region, which tends to be overexposed in Duncan's beautiful photograph* with the Mt. Wilson 100-inch reflector. Duncan has drawn attention to the curved

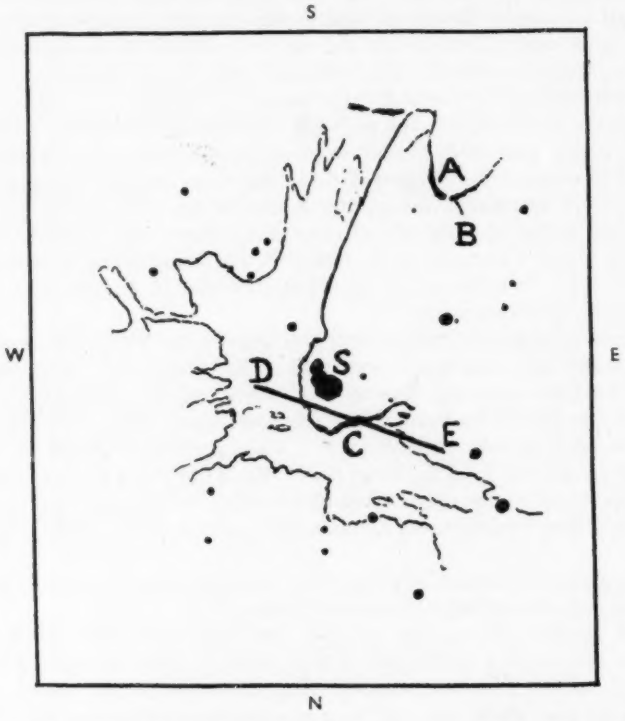


FIG. 4.—*Sketch of M20 showing position of slit (DE) for spectrum of Plate 5 (b); also the bright wisp B at break in rim A (see text).*

bright line of light convex northward (marked A in Fig. 4), i.e. pointing roughly towards and 2'·2 from the central star HD 194492 (S in Fig. 4). This appears to be a typical bright rim to a dark lane, the dark matter, however, being partly veiled by foreground nebulosity. Short exposures with the Radcliffe reflector

* J. C. Duncan, *Ap. J.*, **57**, 142, 1923; and *Mt. Wilson Contr.*, No. 256, 1923.

reveal an extremely interesting feature, which may be of great significance in interpreting the nature of interaction between dark and bright matter at the boundary. Near the northern tip of this rim there is a minute break in the line of light, about 2″·3 of arc across. This might of course be merely due to a narrow belt of absorbing matter lying in front of the rim. However, on the east side of the break there is a faint wisp of bright nebulosity (marked B in Fig. 4), about 16″ long, extending from the bright rim at the break northwards (p.a. about 15°), i.e. on the bright side of the boundary nearer to the central star. This feature will be referred to later in the general discussion.

Short exposures also show that at the boundary between the main dark lane running east and slightly north and the bright nebulosity immediately surrounding the central star S, there is an intensely bright rim about 3″ of arc wide running east and slightly south (marked C in Fig. 4); it also persists faintly round the corner at the junction of the two main dark lanes N.W. of S. This area was selected for spectroscopic study. One spectrum was obtained with slit east–west crossing the 10th mag. companion to S (p.a. 22°, $d=6''$). A second (NS 108) was obtained crossing the bright rim with slit rotated to p.a. 72° instead of the usual 90°; for this spectrum the slit was offset 15″ north of S. The position of the slit is marked DE in Fig. 4. The second spectrum only was run through the microphotometer, as being more suitable for study of the variation of the lines across the bright rim. The length of the analysing slit was this time reduced to correspond to 0·5 mm. on the film. The measures of N1, N2, $H\beta$ and $H\gamma$ are presented in Table III.

TABLE III

Measures of nebular lines across bright rim in NGC 6514 (*Spectrum NS* 108)

x = distance (in seconds of arc) from nearest point to star S along slit, set in p.a. 72°; positive eastwards.

The tabulated quantities are $\log_{10}$ (peak intensity) with arbitrary zero.

The sixth column (mean emission) gives the mean of the preceding four columns.

The seventh column gives the difference between the sums of $H\beta$, $H\gamma$ and N1, N2, and is thus a measure of the relative intensity H/O III.

x	N1	N2	$H\beta$	$H\gamma$	Mean emission	H−N
−22	0·89	0·32:	0·94	0·94	0·77	+0·67
−12	1·14	0·57	1·11	1·04	0·96	+0·44
+ 7	1·18	0·91	1·25	0·97	1·08	+0·13
+11	1·19	1·01	1·24	1·04	1·12	+0·08
+16	**1·35**	**0·39**	**1·36**	**1·27**	**1·09**	**+0·89**
+21	**1·33**	**0·61**	**1·39**	**1·27**	**1·15**	**+0·72**
+26	**1·14**	**0·51**	**1·44**	**1·15**	**1·06**	**+0·94**
+30	0·89	0·20:	1·17	0·92	0·80	+1·00
+35	1·01	0·32	1·10	0·85	0·82	+0·62
+40	0·45:	0·47:	1·11	0·85	0·72:	+1·04:
+45	0·95	0·56	1·14	1·07	0·93	+0·70
+50	1·11	0·60	1·00	1·01	0·93	+0·30
+54	1·05	0·25	1·13	0·75	0·80	+0·58
+64	1·24	0·88	1·28	1·12	1·13	+0·28

There is no direct photograph with suitable calibration available for comparison with the measures on NS 108, but the three rows of figures in heavy type correspond to the regions where the slit was traversing the bright rim as shown on the direct photographs. The slit actually traversed the rim C at about 32°, and assuming the

rim to consist of a straight band 3″ of arc wide, the slit (whose width corresponded to 3″·4 of arc) would transmit a total length of 11″ of arc of rim (apart from guiding errors which would lengthen it further).

From Table III we conclude, regarding this bright rim in the central region of the Trifid:

(1) The nebular lines N1, N2, $H\beta$ and $H\gamma$ rise to a maximum of intensity at or near the rim C.

(2) The ratio of intensity Hydrogen/N1, N2 is greatest at or near the rim; if anything the hydrogen is relatively brightest to the *dark* side of the rim, and N1, N2 certainly maintain an intensity near their maximum on the bright side nearer the central star.

Although the second spectrum (NS 107), with slit east–west, has not been measured with the microphotometer, visual inspection shows that $H\beta$ again reaches maximum intensity relative to N1 to the dark side of the rim west of S. Both spectra show the presence of faint continuous spectrum in the bright nebulosity immediately surrounding S, but its proximity to the bright star rendered it unsafe to attempt photometric measures.

M8 (NGC 6523) (Plate 6 and Fig. 5)

The first direct photograph of this nebula with the Radcliffe reflector (30 minutes on emulsion Kodak 103 A E with red filter) shows the centre completely blacked out. A number of dark areas with rather faint emission rims are shown in the outer regions. A wavy line of nebulosity encircling the southern side of the cluster NGC 6530 is very prominent—much more so than on Duncan's blue photograph.* This probably represents an area of enhancement of hydrogen emission analogous to the bright rims.

No bright rims suitable for spectroscopic study having been found, interest in this nebula was thereafter concentrated on the brightest area west and slightly south of the O star 9 Sgtr. (This star appears in the bottom right corner of Plate 6 but lies outside Fig. 5.) The area is so bright that it has been readily seen in the 74-inch with a full Moon not far away. It was first described by Sir John Herschel as "a kind of elongated nucleus, just following a star (*n*, No. 36).... The proper nucleus is decidedly not stellar, and resembles much more that of the nebula in Andromeda than any other I can call to mind".

Two short exposures (5 mins. and 90 secs.) in the red, and two in the green (Wratten 58 plus Kodak O–F and Kodak 103 AO), obtained with the Radcliffe reflector, effectively isolate this area in $H\alpha$ (+*N* II), continuous, and in *O* III (+continuous) light respectively. The main nucleus east of Herschel 36 is of the shape of an hour-glass, with a dark division 2″ of arc, or less, wide crossing the waist.† Both halves of the hour-glass have an intricate structure with a number of dark divisions. The most remarkable feature, seen best on the 5-min. red exposure (Plate 6(*a*)), is a wisp of bright nebulosity proceeding from the N.W. corner of the hour-glass which curves round and meets the star Herschel 36

* J. C. Duncan, *Ap. J.*, **61**, 4, 1920; and *Mt. Wilson Contr.*, No. 177, 1920.

† Fig. 5 represents a very imperfect attempt to sketch some of the more important features of the area, visible on the original negatives, but largely lost in reproduction. Within the hour-glass the star-like condensations in continuous light of Plate 6 (*b*) are shown. Outside the hour-glass the shadings represent areas of bright $H\alpha$. The area west and north of Herschel 36 is irrelevant to the discussion and has only been roughly sketched. Attention is specially drawn to the $H\alpha$ structure near the stars Herschel 36 and Cordoba 12403, which are both marked in the figure.

in approximately p.a. 0°. It appears to be continued in a fan-shaped extension about 2″ of arc long issuing from the other side of the star in p.a. about 110°. This fan is about 4″ wide at its extremity (as recorded in 5 mins.) and is much narrower close to the star. This apparent association with Herschel 36 might of course be a matter of chance projection. However, there appears to be a second instance; the star Cordoba 12403 (m_{vis} 9·4) seems to mark the termination of a much longer faint wisp which issues from the S.E. corner of the hour-glass. This nebulosity appears to be also linked up with a very bright wisp issuing from the S.W. corner of the hour-glass.

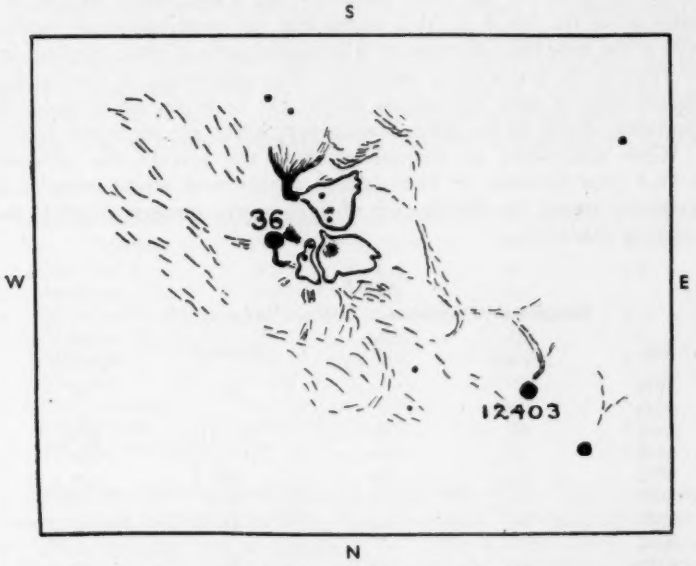


FIG. 5.—*Sketch of M8 nucleus showing Herschel 36 and Cordoba 12403 with associated bright nebulosity. Within the boundaries of the "hour-glass" are shown the star-like condensations revealed in the continuous photograph (Plate 6 (b)). Outside this area are sketched roughly the patterns of bright Hα nebulosity (Plate 6 (a)).*

It is true that both stars Herschel 36 and Cordoba 12403 may be subject to heavy local obscuration, but unless the absorption is of the order of three magnitudes or more, we are faced by two instances of apparent close association of nebula and stars intrinsically fainter than the 6th mag. O star, 9 Sgtr, which might be regarded as the dominating influence in this central part of the nebula.

The shorter of the red exposures shows that in the internal structure of the hour-glass the northern half is brighter on its west side, the southern on its east side. This internal structure appears quite different on the green (mainly continuous) plate (Plate 6 (*b*) and Fig. 5). Each half of the hour-glass shows at least three concentrations which are strongly suspected to be stellar in character, their diameters, with the exception of one in the northern half, are of the order of 1″ of arc, i. e. the size of the smallest star images.

The Kodak 103 A O + W58 combination, which records mainly N1, N2 plus continuous, shows yet another distribution. The difference from the red plates

is especially marked in the southern half of the hour-glass; the S.W. corner near to the exit of the bright wisp, already mentioned, is the brightest part. The suspected faint stars are lost in nebulosity.

Two spectra of this area have been obtained, both with slit east–west. One, exposure 30 minutes, includes the star Herschel 36 and the northern half of the hour-glass; the other, exposure 60 minutes, with the slit about 10″ of arc farther south, refers to the southern half (Plate 6(*c*)). Although the spectra were taken with fairly narrow slit (0·15 mm.), both show a mixture of a typical bright-line nebular spectrum and strong continuous. The intensity of this continuous spectrum relative to N1 has been measured and a calculation similar to that previously given for rim A in M16 shows that the contribution of continuous spectrum to the total light effective on blue photographs is much brighter in this part of M8—of the order of 35 per cent. The appearance of the continuous band in the spectra does not suggest that it can be explained in terms of the condensations, visible on the green photograph, which are suspected to be faint stars. When considered in conjunction with the spectra, the combination Kodak O–F plus Wratten 58 filter (which avoids most strong nebular lines) must certainly record the distribution of light in this nucleus which is mainly continuous in character.

TABLE IV

Emission lines measured in brightest region of M8

λ_{obs}	Int. NS 102	Int. NS 109	Element	Int. Orion
5875	...	1	*He I*	30
5659	? 1n	...	...	...
5007	40	50	[*O* III]	430
4959	12	20	[*O* III]	120
4863	50	80	*Hβ*	100
4472	2	6	*He* I	6
4340	20	25	*Hγ*	35
4200	? 1n	...	...	...
4101	10	15	*Hδ*	20
4069	1n	2	[*S* II], *C* III	2
4026	0·5n	4	*He* I	3
3988	? 0·5	...	...	...
3969	8	10	[*Ne* III], *H*	7+20
3939	...	? 0·5	...	...
3889	6	6	*H*	15
3868	2n	2	[*Ne* III]	20
3853	1	? 1	? *Si* II	2n
3835	3	4	*H*	12
3797	? 1n	2	*H*	10
3727	20	25	[*O* II]	50+30

Table IV lists the mean wave-lengths of emission lines measured on the two spectra, obtained by means of a Hartmann reduction formula, smoothed to the nearest angstrom unit. No comparison spectrum was impressed but a suitable value of n_0 was applied in the formula to give the correct wave-length at *Hγ*. Visual estimates of intensity are listed separately for the two spectra, together with those for the Orion nebula from Wyse's paper.* The chief difference from the Orion nebula is the great weakness of the *Ne* III line 3869 in M8. That this is probably

* A. B. Wyse, *Ap. J.*, **95**, 356, 1942.

only a local difference is shown by the fact that the line is faintly visible in the surrounding nebulosity (see Plate 6(*c*)), but it does *not* strengthen markedly, as do the other emission lines, in the neighbourhood of the hour-glass. *He* 4471 and 4026 appear to be relatively strong in the southern half (NS 109), and another line which is certainly very strong in M8 compared with the Orion nebula is 4069, attributed by Wyse to [*S* II], *C* III. All hydrogen lines appear strong relative to *O* III in M8; *O* III 4363 was not detected, possibly owing to proximity to *Hγ* on the small scale. The general character of the spectra suggests that this part of M8, despite its great brilliance, is one of considerably less ionization than the Orion nebula, and that permitted lines are enhanced relative to the forbidden.

The continuous spectrum has been carefully searched for absorption features. While a number of possible absorptions have been measured, none coincides with a wave-length corresponding to a feature in a likely stellar source, with the possible exception of a shallow depression measured at λ 4304 on NS 109, which may be connected with the G band. For further investigation of this point a dispersion higher than that at present available at the Radcliffe Observatory is imperative.

TABLE V

Relative intensities of nebular lines in selected portions of three diffuse nebulae

Object	N1	N2	*Hβ*	*Hγ*
M16 rim A	14	4·0	10	4·4
M16 near A	17	4·3	10	...
M16 rim B	16	4·2	10	5·0
M16 near B	17	5·6	10	...
M16 rim C	16	4·7	10	5·3
M20 rim	5·3	1·3	10	5·0
M20 near rim	11	3·0	10	6·3
M8 nucleus	12	3·5	10	4·3
M8 near nucleus	19	5·6	10	5·6

2. *Summary of observational data on bright rims.*—The relative intensities of the four principal nebular lines in the blue, as measured in selected portions of the three diffuse nebulae, are summarized in Table V. Hubble*, from slitless spectra, has given for both M16 and M20, N1 + N2(1), *Hβ*(3), *Hγ*(2), and for M8, N1(2), N2(1), *Hβ*(4), *Hγ*(3). The fact that N1 is nearly always brighter than *Hβ* in Table V is probably to be ascribed to the use of an orthochromatic emulsion. The characteristics of the Seed 30 plates used by Hubble are not known to the writer, but probably the sensitivity was falling sharply to the red of *Hβ*. Although the ratio N1/*Hβ* in diffuse nebulae may have been underestimated in the past, due to this cause, it would probably not affect the conclusion reached by various workers that *Hβ* is strong relative to N1 in diffuse compared with planetary nebulae.

Evidence has been brought forward in this paper that *Hβ* is still further strengthened relative to N1, N2 at the bright rims, and this can be seen in summarized form in Table V, especially for M20. Although, as has been mentioned before, the reduced scale on the spectra renders the widths of the rims exceedingly small, it is believed that this conclusion is well established, and further that the continuous spectrum apparent in rim A of M16 provides a negligible contribution to the integrated light. As far as the writer is aware, the only previous attempt to isolate these bright rims for spectroscopic study has

* E. Hubble, *Ap. J.*, 56, 176, 1922.

been made with the McDonald nebular spectrograph, and the negative result for a rim to a dark lane in Taurus was taken as evidence that the rim originated mainly from continuous spectrum.* However, the Taurus lane does not appear to have the sharp boundary presented by those in M16 and M20 and the bright rims may well be different in character.

In addition to the observation that the rims show nebular spectra with strengthened hydrogen lines, the following qualitative data from the appearance on direct photographs are relevant here:

(1) *Curtains* of long, rather straight bright rims are often found as boundaries between dark and bright areas; typical examples, IC 434 and along the whole eastern side of M16 (areas E and F of Fig. 1).

(2) Dark lanes cross the line of these curtains, often continuing for great distances, their edges being marked by bright rims; best example, the lane pointing due west in M16, N.E. of rim A (area G of Fig. 1). (The print for Plate 4(*a*) is too densely exposed to show this beautiful feature.) The whole area between E and B on Fig. 1 represents another case on the grand scale. We refer to this phenomenon as a ***break-through.***

(3) The dark lanes with bright rims usually point towards a bright central star or group of stars and the rim is brightest at the tip nearest to the star.†

(4) Dark lanes sometimes show a phenomenon of ***bifurcation*** at the end where the boundary is sharpest. An excellent example is the most southerly of the three central lanes in M16, the rim being especially bright here in the bay between the two "heads" (H of Fig. 1). Such bifurcation has also been noted (with or without bright rims) in the Eta Carinae nebula.

(5) Finally, the apparent width of these bright rims is of the order of 2″ to 3″ of arc, which, at the accepted distances of these nebulae, corresponds to a linear width of several thousand astronomical units.

3. *Discussion.*—The fact that these rims represent mainly line emission, lends support to the hypothesis that they are areas of excitation at a boundary of interaction between dark and bright matter, similar to that considered by Oort‡ for instance in the shell surrounding Nova Persei. Reflection or diffraction of light from stars hidden by dark lanes is ruled out. There is thus no analogy with the silver lining of terrestrial clouds.

The fact that so many dark lanes show elongation towards a bright star, as already pointed out by Struve, does suggest a direct association of the lanes with the stars and bright nebula, rather than that they happen to be projected against a bright background. If we assume direct interaction, the question of motion relative to the central stars becomes all important. Is the dark matter moving inwards against the bright nebula, or is the nebula being driven away from a central region against a dark obstruction?

The small scale of the Radcliffe spectra eliminates the possibility of detecting anything but very large Doppler shifts.§ An attempt to detect proper motion

* O. Struve and C. T. Elvey, *Ap. J.*, **89**, 119, 1939. (*Note added later.*—Dr van de Hulst has since informed me by letter that he has made very similar and more extensive observations of M16 and other diffuse nebulae at the McDonald Observatory.)

† This fact has been remarked by O. Struve, *Ap. J.*, **85**, 208, 1937, in his list of bright rims, and doubtless by many other observers. Another excellent example near S Mono has been published by W. W. Morgan and F. Shearman, *Ap. J.*, **89**, 554, 1939.

‡ J. H. Oort, *M.N.*, **106**, 159, 1946.

§ A slight distortion in the bright lines in region C of M16 is fully explained by assuming non-uniform illumination of the slit by the rim which was oriented tangentially.

of some features common to the Radcliffe photograph of M16 and Duncan's (R.A.S. Slide No. 432) failed, but at a distance of 2000 parsecs there would appear to be much greater hope of detecting motions spectroscopically. Moreover, in view of the variations of $H\beta/N$, proper motions of nebulosity would probably have to be measured from pairs of plates with very similar colour sensitivities.

In the absence of radial velocities and proper motions, one may fall back on evidence from direct photographs. The phenomenon of *break-throughs* does strongly suggest motion of the dark matter towards the central area. It is extremely difficult to understand how these phenomena could occur in M16 if the motion were in the opposite sense, for in that case one would expect to find instances of dark lanes terminating before the "curtain" is reached. Still stronger evidence of motion towards the central star is afforded by the bright rim A (of Fig. 4) in the Trifid, already mentioned (see Plate 5(*a*), R30, T 35 mm.). Unless the connection of the break in the rim and the faint wisp B to the northern side is an extraordinary coincidence, we must conclude that some force has broken through the rim locally and left a remnant of it on the side nearer to the star in the form of the wisp. It may be concluded as a working hypothesis for M16 and the Trifid that the motion of the oriented dark matter is towards the central stars.

The origin of the continuous spectra detected in these three diffuse nebulae remains a matter of considerable interest. These observations do not suffice to distinguish between reflection and some other process concerned in the interaction between dark and bright matter. But it does seem probable that the "transition type" of nebular spectrum, composed of a mixture of line and continuous spectrum is commoner than has hitherto been indicated.* Moreover the observation of the M8 nucleus, which provides the strongest continuous spectrum, shows that the brightest area of a diffuse nebula may be intimately connected with relatively faint stars—the association being possibly of the T Tauri type. Hubble's statistical law, connecting the extent of diffuse nebulae with the brightness of the central star, refers to the excitation of faint outer nebulosity; but in the brightest area of M8 we seem to be witnessing complex processes in which faint stars, dark and bright matter are all playing important roles. Perhaps the peculiar feature of such nebulae is that while the general environment is determined by an early-type star, interstellar dust is not so strongly repelled by radiation pressure from the fainter stars and can approach to smaller periastral distances.

The writer is deeply indebted to Professor Oort and Dr van de Hulst for reading this paper and offering valuable criticisms.

Radcliffe Observatory,
Pretoria:
1949 *September* 19.

* Since the above was written, my attention has been drawn to similar evidence found by P. Swings and O. Struve, *Ap. J.*, **96**, 310, 1942.

ULTRA-VIOLET EMISSION FROM THE CHROMOSPHERE

R. v. d. R. Woolley and C. W. Allen

(Received 1950 March 8)

Summary

A model of the quiet chromosphere is constructed which gives as good a fit as found possible to data from eclipse observations and to observations of solar radio noise and of the ionosphere, everything being taken at minimum solar disturbance. The model has spherical symmetry and a single value of the kinetic temperature at every height. In the model there is a sharp division between the lower chromosphere, at heights below 6000 km., in which the temperature is 5040 deg., and the upper chromosphere in which the temperature ascends at first very rapidly. It is estimated that the chromosphere emits about 7×10^{14} quanta capable of ionizing terrestrial gases per sq. cm. per sec., of which 6×10^{14} are emitted in line spectra of ions such as *O* VI, *O* V, *N* V, etc, and the remainder in the Lyman continuum of *H*. About half the quanta have an energy greater than 13·6 volts, and are therefore capable of ionizing *O* to O^+ in the ionosphere. The energy is supplied by conduction inwards from the corona.

---

1. *Introduction.*—In a recent paper (1) we calculated the emission spectrum of the corona in the far ultra-violet and drew attention to the part it might play in ionizing terrestrial gases to form the ionosphere. We suggested at that time that an appreciable part of the ionizing radiation received by the Earth's upper atmosphere might come from the chromosphere.

In this paper we attempt to make the same calculations for the chromosphere as we did for the corona : but whereas in the case of the corona values of the kinetic temperature and electron pressure which we felt to be reliable were available, we could not feel satisfied with any account of the temperatures and pressures in the chromosphere which we could find. Since the ultra-violet emission is very sensitive to both temperature and pressure, it has been necessary for us to work out a model chromosphere satisfying as many observational conditions as possible before we could estimate the ultra-violet emission.

2. *Conditions to be satisfied by a model quiet chromosphere.*—While the temperatures and pressures in the reversing layer and corona are reasonably well established, there are difficulties associated with attempts to deduce these quantities in the layer between them, the chromosphere. It has not yet been possible to set up a completely satisfactory system of differential equations representing heat transfer and mechanical equilibrium; moreover, direct observation shows that the actual chromosphere is torn about by disturbances, so that a model exhibiting spherical symmetry may be quite misleading. Nevertheless the idea of a quiet chromosphere with spherical symmetry may be of some use. It is known that the chromosphere, corona and ionosphere all exist at times of zero sunspot activity, and we set out with the idea of a model quiet chromosphere which is intended to represent minimum activity.

A satisfactory model of the quiet chromosphere should fit the following conditions :—

(*a*) The far ultra-violet emission should equal the amount required to produce the minimum (zero sunspot) ionosphere.

(*b*) This emission should equal the amount of energy conducted inwards from the corona.

(*c*) The emission of thermal radio noise should fit quiet day observations on various wave-lengths.

(*d*) The density scale height of the low chromosphere should fit eclipse observations of decrement of the intensity of low excitation lines with height.

(*e*) The density scale height should fit eclipse observations of the decrement of strong Balmer lines extending to considerable heights.

(*f*) Electron densities should agree with eclipse values (1) from the intensity of the Balmer continuum and (2) from the resolution of Balmer lines.

(*g*) Electron temperatures should agree with eclipse observations of the Balmer continuum.

(*h*) Kinetic temperature of atoms should agree with eclipse observations of line width.

(*i*) Eclipse observations show that both high and low excitation lines (e. g. *He* and *Fe*) appear to originate at the same level. The model should explain this.

(*j*) The density scale height should be not less than the gravitational scale height for the kinetic temperature at the level. If it is greater, there should be some reason for it.

(*k*) The pressure should not increase outwards.

We have not found it possible to satisfy all these conditions with a spherically symmetrical model possessing a definite single value of the temperature at each height. We have however been unwilling to introduce a parameter representing variations of temperature from one part of the chromosphere to another at the same level, largely because this would introduce too much freedom into the model. Instead the model uses a mean temperature for each height.

It should be said here, briefly, that we have not taken (*e*) into account, because of the difficulty of calculating populations of the lower hydrogen excited states; that we have been unable to satisfy (*h*), which appears irreconcilable with (*a*); and that we cannot satisfy (*i*), which seems to demand variations of temperature at the same height.

3. *Division of the chromosphere into two regions.*—It is convenient to define three definite levels in the chromosphere :—

(*a*) The base of the chromosphere. Quantities evaluated for this level will be given the subscript *o*.

(*b*) The top of the chromosphere, which will be a specified level in the corona (subscript *c*).

(*c*) The level where hydrogen is one-half ionized (subscript *i*).

The region between (*a*) and (*c*) we call the low chromosphere. The flash spectrum is emitted mainly from the low chromosphere, in fact most of the phenomena generally regarded as chromospheric occur in this region. On the other hand, the region between (*c*) and (*b*), the upper chromosphere, is almost entirely coronal in character. From it come the radio noise and most of the far ultra-violet emission.

Our reasons for supposing that the division between the two levels is sharp are as follows :—

(1) The equation for the conduction of heat inwards from the corona, which leads to a temperature distribution in good agreement with radio noise observations, demands a very sharp drop in temperature at the bottom of the region controlled by conduction.

(2) A layer of extensive thickness at any temperature sufficient to ionize hydrogen appreciably leads to far too great an emission of Lyman quanta.

(3) The corona only escapes from radiative control, which would pin it to the Sun's surface temperature, because it is optically thin. A few hundred kilometres of neutral hydrogen (at a density appropriate to the middle chromosphere) destroy the condition of optical thinness, through the heavy absorption in the Lyman continuum. Wherever hydrogen is substantially neutral, the material cannot be far from radiative equilibrium and the temperature cannot be far from the photospheric value.

4. *The base of the chromosphere.*—For the base of the chromosphere it is reasonable to assume equilibrium conditions at the Sun's surface temperature. The density scale height (i. e. the height within which the density falls one exponential factor) will then be

$$H = \frac{kT}{\mu g} = 1{\cdot}2 \times 10^7 \text{ cm.},$$

where k is the Boltzmann constant, T the temperature for which we adopt 5040 deg. or $\Theta = 5040/T = 1{\cdot}0$, μ the mean atomic mass for which we adopt $2{\cdot}1 \times 10^{-24}$ g., and g gravity $2{\cdot}74 \times 10^4$ cm. sec.$^{-2}$.

At this level metals will be almost fully ionized and will provide most of the electrons. If the hydrogen : metal number ratio is 6000 : 1 (**1**), we obtain

$$\log (N_e)_0 = \log (N_H)_0 - 3{\cdot}8, \qquad (4.1)$$

where N_e and N_H are the numbers of electrons and neutral hydrogen atoms per cm.^{3}, and $(\ )_0$ signifies the base of the chromosphere.

By definition, see e. g. Wildt (**2**), the continuous absorption for a transverse beam through the base of the chromosphere integrates to unity. For electron densities of the order of 10^{12} and $\Theta = 1$ the absorption is mostly by negative hydrogen ions (**3**) and is given as $9{\cdot}32 \times 10^{-26}$ per H atom and per unit electron pressure. The absorption coefficient is then

$$6{\cdot}5 \times 10^{-38} N_H N_e \text{ cm.}^{-1}.$$

The integrated absorption through a transverse beam becomes

$$6{\cdot}5 \times 10^{-38} (N_H N_e)_0 (\tfrac{1}{2} H 2\pi R_\odot)^{1/2},$$

where $R_\odot$ is the Sun's radius, and $\frac{1}{2}H$ is the scale height of the product $N_H N_e$. Equating this to unity and using (4.1) gives

$$\log (N_H)_0 = 15{\cdot}9, \qquad \log (N_e)_0 = 12{\cdot}1,$$

and from Saha's equation the proton concentration N_p is

$$\log (N_p)_0 = 11{\cdot}2.$$

These values of $(N_H)_0$ and $(N_e)_0$ are a little higher than those determined by Wildt (**2**), the reason being that we have assumed the scale height of gravitational

equilibrium, whereas Wildt extrapolates the observed scale height to the base of the chromosphere.

5. *The top of the chromosphere.*—It is necessary to choose a particular level in the corona to represent the top of the chromosphere at a height h_c. Baumbach (**4**) has given the relation between electron density and height, shown in Table I.

TABLE I

The Lower Corona

Height	h' (unit $=10^9$ cm.)	0	2	4	6	8	10
Electron density	N_e (unit $=10^8$ cm.$^{-3}$)	4·5	3·3	2·4	1·8	1·3	1·0
$h'N_e^2$	(unit $=10^{26}$ cm.$^{-5}$)	0	2·2	2·3	1·9	1·4	1·1

The ultra-violet and the radio emission depend on $(h'N_e^2)_c$, and it is seen from Table I that $h'N_e^2$ is fairly independent of h'. Hence the particular choice is not important. We assume Baumbach's heights are measured from h_i, the height of the half-ionized layer, and fit the chromosphere to the corona at the height $h_c = h_i + 4 \times 10^9$ cm., which is roughly the height where coronal lines reach their greatest intensity (**5**). Choosing 10^6 deg. K. as the corona temperature (**1, 6**) we have the adopted values for the top of the chromosphere as follows:

$$N_c \equiv (N_e)_c = (N_p)_c = 2{\cdot}4 \times 10^8 \text{ cm.}^{-3},$$

$$T_c = 10^6 \text{ deg. K.},$$

$$h_c - h_i = 4 \times 10^9 \text{ cm.},$$

$$(h_c - h_i)N_c^2 = 2{\cdot}3 \times 10^{26} \text{ cm.}^{-5}.$$

The electron densities of the corona have been analysed more recently by van de Hulst (**7**), who finds a considerable change with the sunspot cycle. His values for sunspot minimum give $(h_c - h_i)N_c^2 = 3 \times 10^{25}$, and would cause a considerable change in the values calculated in this paper. However, we prefer to accept Baumbach's data as a standard and modify the results if found necessary.

6. *The level where hydrogen is half ionized.*—We suppose that the character of the chromosphere changes very rapidly at the level where hydrogen is half ionized. Above this level there are few hydrogen atoms and the chromosphere is optically thin; but below this level the strong Lyman continuum makes the optical depth rise rapidly. The absorption of the H atom beyond the Lyman limit is approximately 10^{-17} cm.^{2}, so that, if the neutral hydrogen density N_H were only 10^{10}, unit optical thickness in the continuum would be reached in 100 km.

A lower limit to the atom and electron populations at this level may easily be obtained. Although gravity can be reduced by supporting forces (such as radiation pressure) the net resultant force on the materials must be inwards, otherwise the chromosphere would disperse outwards; hence the pressure cannot increase outwards. We must have, therefore,

$$\left.\begin{aligned} &(N_H + N_p + N_e)_i T_i \geqslant (N_p + N_e)_c T_c, \\ \text{or} \quad &3N_i T_i \geqslant 2N_c T_c, \end{aligned}\right\} \qquad (6.1)$$

the subscript p referring to protons, N_i is the number of protons, electrons, or neutral H atoms per cm.^{3} at the half-ionized level, and N_c the number of protons or electrons per cm.^{3} at the top of the chromosphere.

After carrying out a number of calculations upon models in which the pressure was allowed to decrease with height, and in which the amount of ultra-violet radiation given by the model was always too high (i.e. predicted too great an ionosphere), the authors felt forced to the conclusion that the decrease of pressure with height in the high chromosphere is very small: the pressure has therefore been taken to be constant, that is,

$$3N_iT_i=2N_eT_e. \tag{6.2}$$

As the half-ionized level lies immediately above a layer of unit optical depth in the Lyman continuum and is subjected to some ultra-violet radiation from above, the dilution factor of the Lyman continuum radiation must be close to unity. It can be shown that the effect of collisions upon the ionization balance is small, so that the degree of ionization is given well by Saha's equation: this gives a second relation between $(N_e)_i$ and T_i. From Saha's equation and (6.2) we find

$$N_i\equiv(N_e)_i=(N_p)_i=(N_H)_i=2{\cdot}5\times10^{10}\,\text{cm.}^{-3},$$
$$T_i=6300\,\text{deg.}$$

7. *The lower chromosphere.*—By lower chromosphere we mean the region between the photosphere and the level where hydrogen is half ionized. Since the lower chromosphere is optically thick just beyond the Lyman limit, hydrogen ionization is adjusted to the Sun's surface temperature, and we expect the kinetic temperature to be close to this temperature.

There is ample evidence that the scale height H in most of this region is considerably greater than the gravitational value $1{\cdot}2\times10^7$ cm. For example, Cillié and Menzel (**8**) give about 6×10^7 cm. for metals at about 1100 km. height, and Wildt (**2**) gives $2{\cdot}4\times10^7$ cm. at a height of 800 km. All observations show H to increase with height h. Since we have found the same hydrogen to metal ratio in the corona as obtains in the reversing layer (**1**), we may assume the scale height to be the same for all atoms (provided that all states of ionization are taken together). We adopt

$$H=1{\cdot}2\times10^7+0{\cdot}20h,$$

where the heights are in centimetres. This formula is in reasonable agreement with eclipse observations and reduces to the gravitational value when $h=0$. As we suppose that the kinetic temperature does not increase much with h in the *lower* chromosphere, we must attribute most of the rise in scale height to turbulence of the kind suggested by McCrea (**9**).

The relation between $N(=N_H+N_p)$ and h is

$$\frac{N}{N_0}=\left(1+\frac{0{\cdot}2h}{1{\cdot}2\times10^7}\right)^{-1/0{\cdot}2}.$$

From this formula, and from $\log N_0=15{\cdot}9$, $N(=N_H+N_p)$ reaches 5×10^{10} where $h=6\times10^6$ cm., or $h_i=6\times10^8$ cm.

In Table II, where we tabulate various quantities for the entire chromosphere, we suppose T to remain at 5040 deg. up to $h=5\times10^8$ cm., and to rise to 6300 deg. at $h=6\times10^8$ cm. The values of N_e and N_p in the lower chromosphere are calculated from Saha's equation.

8. *The upper chromosphere.*—The temperature distribution of the upper chromosphere should be governed mainly by the conduction of energy

TABLE II

Model Chromosphere

Height ($\times 10^8$ cm.)	T	log N	log N_e	log N_p	log N_H
0	5040	15·9	12·1	11·2	15·9
1	5040	13·8	10·6	10·6	13·8
2	5040	12·7	10·0	10·0	12·7
3	5040	12·0	9·7	9·7	12·0
4	5040	11·5	9·4	9·4	11·5
5	5040	11·0	9·2	9·2	11·0
6	6300	10·7	10·4	10·4	10·4
6·001	14500	10·2	10·2	10·2	9·4
6·01	36000	9·8	9·8	9·8	5·7
6·1	91000	9·4	9·4	9·4	3·9
7	230000	9·0	9·0	9·0	2·6
8	300000	8·9	8·9	8·9	2·3
10	400000	8·8	8·8	8·8	2·1
14	526000	8·7	8·7	8·7	1·8
19	645000	8·6	8·6	8·6	1·6
26	760000	8·5	8·5	8·5	1·4
33	850000	8·5	8·5	8·5	1·3
39	930000	8·4	8·4	8·4	1·2
46	1000000	8·4	8·4	8·4	1·1

N=number of atoms and ions per cm.^{3}, N_e=electrons per cm.^{3}, N_p=protons per cm.^{3}, N_H=neutral H atoms per cm.^{3}, T=kinetic temperature in degrees K.

inwards (10). Chapman and Cowling (11) have shown that the thermal conductivity of an electron gas is given by

$$\lambda_1 = \frac{75}{16}\left(\frac{k}{\pi m}\right)^{1/2}\frac{k^3}{e^4}\,T^{5/2}/A_2(2),$$

where $A_2(2)$ is $9{\cdot}2\log(4kT/N_e^{1/3}e^2)-2$ or about 50 for the chromosphere, and the other symbols have their usual meaning. For a mixture of protons and electrons in equal numbers the conductivity will be half as great as this, since for a given electron density the free path of electrons will be halved. The heat conductivity in the upper chromosphere then becomes

$$\lambda = \eta T^{5/2},$$

where

$$\eta = 0{\cdot}5 \times 10^{-6}.$$

Then the equation of heat conduction is

$$\epsilon = \eta \cos\theta\, T^{5/2}\frac{dT}{dh}, \tag{8.1}$$

where ϵ is the rate of heat conduction inwards, and θ is the angle between the magnetic field and the vertical. The $\cos\theta$ term is introduced because in the rarefied chromosphere the ions will have freedom of movement only in the direction of the magnetic field. The amount of heat conducted will decrease inwards because some of the energy will be radiated away. To allow for this decrease we adopt the relation

$$\epsilon = \epsilon_c\left(\frac{h-h_i}{h_c-h_i}\right)^n = \epsilon_c y^n, \tag{8.2}$$

where y is the height above h_i in terms of the height of the upper chromosphere $(h_c - h_i)$ and n has yet to be determined. Equations (8.1) and (8.2) combine to give

$$T^{5/2}\frac{dT}{dy} = \frac{\epsilon_c(h_c-h_i)}{\eta\cos\theta}\cdot y^n, \tag{8.3}$$

with the boundary conditions $T = T_i$ at $y = 0$ and $T = T_c$ at $y = 1$, the extra condition imposing a relation between the constants of the differential equation.

If we neglect T_i/T_c compared with unity the solution of (8.3) is

$$(T/T_c)^{7/2} = y^{n+1}, \tag{8.4}$$

and the relation between the constants is

$$\epsilon_c = \frac{2(n+1)\eta\cos\theta}{7(h_c-h_i)}T_c^{7/2}. \tag{8.5}$$

The relation (8.5) gives ϵ_c the amount of heat conducted inwards from the corona. Taking the mean value of $\cos\theta$ to be $\frac{1}{2}$ and using $n = \frac{2}{5}$ to fit the radio observations as explained below, and with $(h_c - h_i) = 4 \times 10^9$ cm., $T_c = 10^6$ (Section 5), we get from (8.5)

$$\epsilon_c = 1{\cdot}7 \times 10^4 \text{ erg cm.}^{-2}\text{ sec.}^{-1}.$$

This is less than estimates given by Alfvén (**12**) ($2{\cdot}2 \times 10^5$) and Giovanelli (**10**) (6×10^5).

The temperature distribution being given by (8.4), the distribution of the electron population follows from the assumption that the pressure is constant (as in (5.2)), for then

$$N_e = N_e T_c/T = N_c y^{-(2n+2)/7}, \tag{8.6}$$

or with $n = \frac{2}{5}$,

$$\frac{N_c}{N_e} = \frac{T}{T_c} = y^{2/5}.$$

It will be remembered that $N_c = (N_e)_c$. Values of N_e and T as a function of height from this relation are given in Table II. In this table ionization of hydrogen in the *upper* chromosphere is found from the collision ionization formula (**1**),

$$\frac{x}{1-x} = \frac{S}{\alpha},$$

where x is the fraction of ionized atoms, S the ionization collision coefficient (for hydrogen $4{\cdot}3 \times 10^{-10}T^{1/2}\exp(-11600\chi/T)$, with ionization potential χ in volts), and α the recombination coefficient (for hydrogen $1{\cdot}8 \times 10^{-11}T^{-1/2}$ in the chromosphere).

We have now to justify our choice of n. This has been taken to give the best fit of quiet day solar noise intensities on different wave-lengths. The effective temperature at the centre of the disk T_E is related to the kinetic temperature T at various levels by

$$T_E = \int_0^\infty Te^{-\tau}\,d\tau, \tag{8.7}$$

where τ is the radio optical depth, equal to $\int_h^\infty \kappa\,dh$ at height h, κ being the absorption coefficient per cm. Martyn (**13**) has given a convenient expression for

this (which can be deduced either from the Kramers-Gaunt formulae or from the Lorentz-Chapman formulae), namely,

$$\kappa = \frac{4}{3}\left(\frac{2}{\pi}\right)^{1/2} \frac{e^6}{c(mk)^{3/2}} \frac{N_e^2}{f^2 T^{3/2}} \ln(3kT/2hf)$$

$$= 1{\cdot}08 \times 10^{-23} N_e^2 \lambda^2 T^{-3/2} \ln(3kT\lambda/2hc).$$

Using the value 13 for the slowly varying term this reduces to

$$\kappa = aN_e^2 \lambda^2 T^{-3/2},$$

where $a = 1{\cdot}4 \times 10^{-22}$.

Now

$$\tau = \int_h^\infty \kappa \, dh = a\lambda^2 \int_h^\infty N_e^2 T^{-3/2} \, dh = a\lambda^2 N_c^2 T_c^2 (h_c - h_i) \int_y^1 T^{-7/2} \, dy,$$

and since $T^{7/2} = T_c^{7/2} \cdot y^{n+1}$ the integration can be easily carried out, giving

$$\tau = (h_c - h_i) N_c^2 a T_c^{-3/2} \lambda^2 (y^{-n} - 1) n^{-1}.$$

If we now put $m = (2n+2)/7n$, $C = (h_c - h_i) N_c^2 a T^{-3/2}$ $(= 3{\cdot}2 \times 10^{-5})$ and $z = C\lambda^2/n$, then

$$\tau = z(y^{-n} - 1),$$

from which

$$y^n = \frac{z}{z+\tau}.$$

Now by (8.4) $T/T_c = y^{2(n+1)/7} = y^{mn}$, so that $T/T_c = (z/z+\tau)^m$. We can now integrate (8.7), the result being

$$T_E = T_c z^m \int_0^\infty (z+\tau)^{-n} e^{-\tau} \, d\tau$$

$$= T_c z^m e^z Ei_m(z).$$

Here

$$Ei_1(z) = Ei(z) = \int_z^\infty x^{-1} e^{-x} \, dx,$$

$$Ei_2(z) = \frac{1}{z} e^{-z} - Ei_1(z),$$

$$2Ei_3(z) = \frac{1}{z^2} e^{-z} - Ei_2(z),$$

$$(m-1)! \; Ei_m(z) = z^{-(m-1)} e^{-z} - Ei_{m-1}(z).$$

(The function $Ei(z)$ is given by Jahnke and Emde (**14**) and called $-Ei(-z)$. It is much more convenient to tabulate $e^z Ei(z)$ which varies comparatively slowly with z. This has been done in typescript tables used in the present work. Where m is not a whole number, the Incomplete Γ function can be used.)

For $n = 0$, m becomes infinite and the solution of (8.7) is

$$T_E = T_c \frac{7C\lambda^2}{7C\lambda^2 + 2}.$$

For the case which we actually adopted, $n = \frac{1}{5}$ and $m = 1$, giving

$$T_E = T_c z e^z Ei(z).$$

The calculated values of T_E for $n=0$, $m=\infty$; $n=\frac{1}{2}$, $m=2$; and $n=\frac{2}{3}$, $m=1$ are shown in Fig. 1. In the same diagram we show various estimates of T_E obtained from observations. Original observations give values of T_A, the apparent temperature of the whole Sun, which as a result of limb brightening and corona spreading is greater than T_E. Recent calculations by Nicolet (**15**) can be interpreted to show that $T_E/T_A \simeq 0{\cdot}5$ and is almost independent of frequency in the

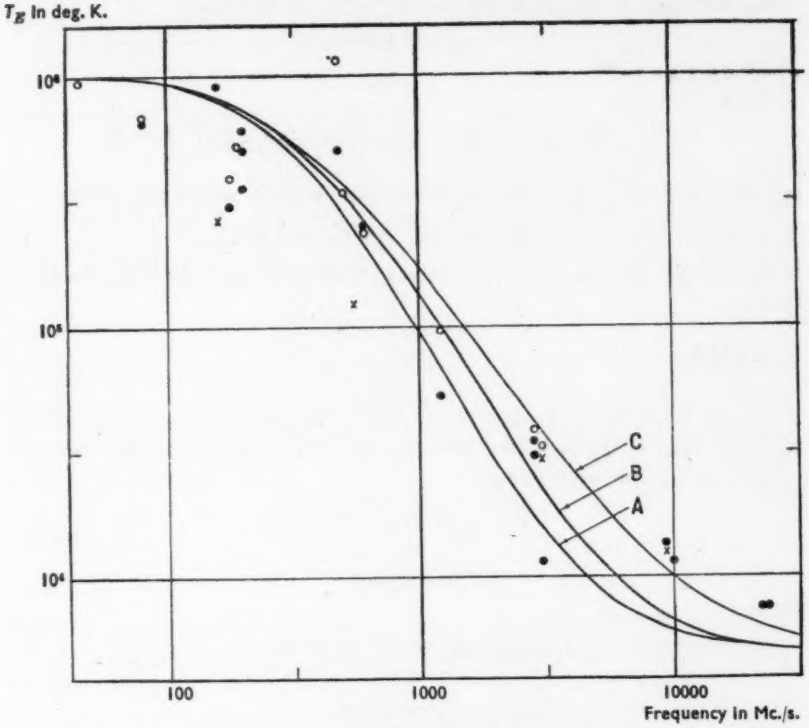


FIG. 1.—*Relation between frequency and effective temperature of the centre of the Sun's disk.*

A: *Calculated* $n=0$, $m=\infty$.
B: *Calculated* $n=\frac{1}{2}$, $m=2$.
C: *Calculated* $n=\frac{2}{3}$, $m=1$.
●: *Observations listed by Pawsey and Yabsley.*
○: *Observations from Quarterly Bulletin.*
×: *Other observations.*

range 100 to 3000 Mc./s. Eclipse observations (**16**) on 600 Mc./s. yield a similar factor. However, for the shortest wave-lengths we must expect

$$T_E \simeq T_A \simeq 5000 \text{ deg. K.},$$

hence we obtain an estimate of T_E from the formula

$$T_E = \tfrac{1}{2}(T_A + 5000 \text{ deg.}).$$

Values of T_A for quiet Sun are obtained from (*a*) a list prepared by Pawsey and Yabsley (**17**) (containing observations by various authorities); (*b*) low values from

daily observations from the *Bulletin of Solar Activity* 1949 (in preparation; containing observations from the Cavendish Laboratory, Cambridge; National Bureau of Standards, Washington; National Research Council, Ottawa; Radiophysics Laboratory, Sydney; and Commonwealth Observatory, Canberra); and (*c*) recent observations by Minnett and Labrum (**18**); Piddington and Hindman (**19**); Laffineur and Houtgast (**20**); and the S.P.I.M. (**21**).

In Fig. 1 it is seen that the higher frequencies show best agreement for $n = \frac{2}{3}$ and $m = 1$. It might be noticed also that on the lower frequencies better agreement would have been obtained if a coronal temperature $T_c = 700{,}000$ deg. had been adopted instead of 1,000,000 deg.

A chromospheric model based on solar noise intensities has been devised by Piddington (**22**). His analysis is similar in intention to ours but makes very different use of the optical evidence. His chromosphere does not have the sharp dividing level. The fact that both his chromosphere and ours are in agreement with the solar noise observations shows that these observations alone are not sufficient to define the chromospheric model.

9. *Discussion of the model.*—Several attempts have been made to interpret the various observations and set up a model of chromospheric conditions. These of course differ in detail but they fall into two main classes, according to the kinetic temperature given to the optically observable chromosphere, the lower chromosphere of our model. There are two distinct views :

(*a*) that the temperature of the optical chromosphere is about 30,000 deg., and
(*b*) that it is about 5000 deg.

The main arguments on each side are

(*a*) for high temperatures, arguments from

(i) thermal line width measurements (**23**),
(ii) scale height measurements (**2**), (**8**),
(iii) the emission of *He* lines;

and (*b*) for the low temperature, arguments from

(i) radio noise measurements (**22**).
(ii) excitation temperature measurements (**24**),
(iii) solar ultra-violet emission as indicated by the ionosphere,
(iv) electron temperature (**25**),
(v) the non-appearance of forbidden lines (**26**).

In making a decision we were mainly influenced by (*b*) (iii), as we were forced to adopt the low temperature to bring the Lyman continuum emission down to a level allowed by the ionosphere. On the other hand, a high temperature chromosphere has been chosen by Giovanelli (**27**) and by Thomas (**28**).

Once the temperature of the low chromosphere has been decided, the model only depends on the base conditions and scale height. If we compare our model with requirements listed in Section 2 we find (*a*), (*f*) and (*g*) in fair agreement with eclipse observations (**2**), (**8**), (**29**), but (*h*) and (*i*) in disagreement. The model offers no reason why the scale height should be greater than the gravitational scale height, but McCrea's suggestion of turbulent motion (**9**) may be accepted. It is possible that the excitation of *He* and He^+ lines (*i*) may be due to a coincidence of a *He* wave-length with a far ultra-violet coronal emission : van Dijke suggests (**30**) that the $2s^3S$ metastable state of *He* is somehow excited, and this could be due to

resonance with *Mg* X radiation from the corona. The wave numbers concerned (**31**) are *He* I 159850 and *Mg* X 159929. Since the latter spectrum line has not actually been observed, the coincidence may be closer than this value suggests.

The failures encountered by our model might be surmounted if we abandoned spherical symmetry and postulated a chromosphere made up of pockets of material at very different temperatures. This is of course strongly suggested by the irregular appearance of the chromosphere and by the apparent presence of high and low excitation lines at the same level. We have however, as has been said before, been unwilling to introduce more variables into our model.

A feature of our model is the sharp dividing level between upper and lower chromospheres, already advocated on very different grounds by Wurm (**26**). In the real chromosphere this division cannot occur everywhere at the same height without distortion. Such distortion may not have much effect on the amounts of far ultra-violet and microwave energy emitted by the centre of the disk, though there may be some reduction of the microwave limb brightening. The emission of ultra-violet radiation *is* however very greatly influenced by the thinness of the skin separating lower and upper chromospheres. Any smoothing of the (mean) temperature curve at the dividing point greatly increases the emission in the Lyman continuum and other spectra.

10. *Ultra-violet emission.*—The study of the chromosphere just described was initiated for the prime purpose of determining the emission of ultra-violet radiation. An approximate estimate of the number of ultra-violet quanta emitted from the Sun's surface at sunspot minimum may be obtained from ionospheric studies (**32**), (**33**), giving for the E region 2×10^{13} quanta cm.$^{-2}$ sec.$^{-1}$, and for both the F_1 and F_2 region 10^{14} quanta cm.$^{-2}$ sec.$^{-1}$. The wave-lengths of these quanta are not known but there is good reason to suppose that they are below the Lyman limit at 912 A., and we give our attention to these wave-lengths, i. e. to excitation energies greater than 13·6 volts. All early attempts to construct a model chromosphere gave far too many quanta and this has been taken into consideration in choosing the model we have described.

The black-body emission above 13·6 volts from the low chromosphere at 5040 deg. is 2×10^{11} quanta cm.$^{-2}$ sec.$^{-1}$ and negligible. However, near the half-ionized level the temperature has risen sharply to 6300 deg., for which the black-body emission is $1{\cdot}4 \times 10^{14}$ quanta cm.$^{-2}$ sec.$^{-1}$. On account of the rapid temperature change the low chromosphere emission will be less than this, probably about half, but it will be of the same order as the emission required to produce one of the ionospheric layers.

The upper chromosphere is optically thin in the Lyman continuum, so that the quanta emitted can be determined simply by counting the recombinations. If α is the recombination coefficient, the total emission per unit time from a column of unit cross-section is

$$Q = \int_{h_i}^{h_c} N_e N_p \alpha \, dh = \int_{h_i}^{h_c} N_e^2 \alpha \, dh.$$

Putting $\alpha = bT^{-1/2}$ (where $b = 1{\cdot}8 \times 10^{-11}$)

we have generally

$$Q = (h_c - h_i) N_c^2 b T_c^{-1/2} \int_0^1 y^{-5(n+1)/7} \, dy$$

$$= \frac{7b(h_c - h_i) N_c^2 T_c^{-1/2}}{2 - 5n},$$

but this formula breaks down if $n \geqslant \frac{3}{5}$, in which case it is necessary to set the lower boundary not at $y=0$ but an appropriate value $y=y_1$, giving the lower temperature $T_i=6300$ deg. instead of T_i zero. For $n=\frac{3}{5}$,

$$Q=(h_e-h_i)N_e^2bT_e^{-1/2}\ln(1/y_1),$$

and $T_i=6300$ deg. gives $1/y_1=316{,}000$. Then

$$Q=5{\cdot}2\times10^{13}\text{ quanta cm.}^{-2}\text{ sec.}^{-1}.$$

These quanta are emitted in all directions, of which half are outwards, so that the outward emission is

$$2{\cdot}6\times10^{13}\text{ cm.}^{-2}\text{ sec.}^{-1},$$

which is not great enough to influence materially the total Lyman emission from upper and lower chromosphere : we have already estimated about 10^{14} cm.$^{-2}$ sec.$^{-1}$ from the lower chromosphere.

It should be noticed that this result has only been achieved by admitting the sharp rise of temperature above the half-ionized layer. A chromosphere at 30,000 deg. would give far too much radiation to be allowed by the ionosphere, for, if the chromosphere were neutral, it would give 30,000 deg. black-body radiation in the Lyman continuum, which is too intense; and if it were ionized it would have to have a sharp temperature boundary at its base to keep the recombination radiation down. The boundary in our model is given by the conductivity equation, but there seems no reason for another sharp boundary at a lower level below a 30,000 deg. plateau as implied by Giovanelli (**27**).

In order to compute the line emission from the upper chromosphere we use the expression derived in our earlier paper (**1**). The number of quanta emitted per unit volume and time is taken to be equal to the number of atoms excited to the upper state by electron collisions, that is equal to

$$N_aN_e\frac{12\pi e^4f}{(2\pi mkT)^{1/2}}\frac{1}{\chi'}\{\exp(-y)-yEi(y)\}, \qquad (10.1)$$

where N_a is the concentration of atoms in the particular state of ionization being considered, χ' is the excitation potential (in ergs) of the strongest transitions, f the oscillator strength, and $y=\chi'/kT$. To simplify the calculations we assume that each atom is in one particular state of ionization in a certain range of chromospheric height. The level dividing each ion from the next higher ion is obtained by putting $x/(1-x)=1$, where x is the fraction ionized. We have the relation

$$\frac{x}{1-x}=\frac{S}{\alpha},$$

where S is the collision ionization coefficient and α the recombination coefficient. We use values that are generalized from our earlier paper as follows:

$$S=3\times10^{-8}T^{1/2}\chi^{-2}\exp(-11600\chi/T) \qquad \text{(with } \chi \text{ in volts)}$$

$$\alpha=1{\cdot}5\times10^{-11}Z^2T^{-1/2},$$

from which we find that the temperature of the dividing level is given by the solution of

$$T=5\times10^{-4}Z^2\chi^2\exp(11600\chi/T) \qquad \text{(with } \chi \text{ in volts).}$$

From the model, the degree of ionization Z, and the ionization potential χ, it is possible to determine the height range occupied by each ion. From (10.1) we determine the emission per unit volume at the top and the bottom of the range.

TABLE III

	Ion	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV
He	χ' log Q'	41 12·0													
C	χ' log Q'	9 11·3	13 12·1	8 13·6	304 8·5	367 9·6									
N	χ' log Q'	11 11·7	13 12·6	16 13·1	10 13·9	430 8·7	500 7·5								
O	χ' log Q'	15 11·6	15 13·1	16 13·5	20 13·8	12 14·3	573 8·3	653							
Ne	χ' log Q'	27 9·5	25 11·6	23 12·5	22 13·1	22 13·1	25 13·3	16 13·5							
Mg	χ' log Q'	4 11·7	53 8·8	39 11·6	35 12·3	31 12·8	29 12·9	29 13·0	34 13·1	20 13·4	1352				
Si	χ' log Q'	7	10 11·2	9 12·2	104 9·5	50 12·0	45 12·2	34 12·7	36 12·7	36 13·2					
Fe	χ' log Q'	5	11 11·5	22 11·3	32 11·4	42 11·4	53 11·1	64 11·1	70 11·9	36 12·4	34 12·7	34 12·4	37 12·6	38 12·5	42 12·9

Excitation potentials in volts χ', and rate of quantum line emission per cm.^{2} of chromosphere per sec. Q'.

If these agree within about a factor of 3 we take a mean and multiply by the height to obtain the total emission. If they differ we use the difference to determine a scale height of the quantity and multiply this by the higher value. In order to apply the procedure we require values of the element/hydrogen ratios, ionization potentials χ, and excitation potentials χ' and oscillator strengths f for the strongest lines. The element/hydrogen ratios are from our earlier paper (**1**), and both excitation and ionization potentials from Mrs Moore Sitterly's recent tables (**34**). It was assumed that most of the energy in each spectrum was emitted in the transition from the lowest configuration to the next configuration of opposite parity. For neutral atoms it would be appropriate to take $f=1$, but when the two configurations have the same total quantum number, as is usually the case, the higher ions become more hydrogen-like and f tends to zero for this transition (**35**). Thus for *Mg* X the f-values, which may be determined from the tabulation of Bates and Damgaard (**36**), are as follows :—

Transition	$2s-2p$	$f=0{\cdot}12$,
	$2s-3p$	$f=0{\cdot}33$,
	$2s-4p$	$f=0{\cdot}08$.

We adopt the value of 0·2 for f throughout, and the emission is divided by 2 to give Q' the outward flux per cm.^{2}. Values obtained for the more abundant ions are shown in Table III. The total outward emission is about 3×10^{14} quanta cm.$^{-2}$ sec.$^{-1}$ for radiation above 13·6 volts (and therefore capable of ionizing O to O^+ in the Earth's atmosphere) and another 3×10^{14} quanta between 10 and 13·6 volts. These values are reasonably close to the requirements of the F regions of the ionosphere, namely 10^{14} quanta cm.$^{-2}$ sec.$^{-1}$. According to our calculations, line emission from the chromosphere is the most important source of solar far ultra-violet emission, the chromosphere being more important than the corona.

We find that both continuous Lyman emission and emission in the lines of ionized atoms are of the right order of magnitude to produce an ionospheric layer; but our calculations are not accurate enough to justify an attempt to identify any particular ionospheric layer with a particular solar emission.

It may be noticed that the calculated solar emission, which is rather high for the ionospheric requirement, would be reduced if we were to use van de Hulst's electron densities (**7**) for the corona, instead of Baumbach's (**4**).

It remains to be shown that the calculated emission of quanta does not require more energy than is conducted inwards from the corona. The various emissions considered have an average energy of about 13 volts, or 2×10^{11} ergs, so that the emission of about 7×10^{14} quanta requires 14,000 ergs cm.$^{-2}$ sec.$^{-1}$, compared with 17,000 ergs cm.$^{-2}$ sec.$^{-1}$ conducted inwards according to equation (8.5). The model is therefore consistent in this respect, and it is unnecessary to postulate a source of energy in the chromosphere as distinct from the corona.

Commonwealth Observatory,
Mount Stromlo, Canberra:
1950 *February* 28.

References

(**1**) R. v. d. R. Woolley and C. W. Allen, *M.N.*, **108**, 292, 1948.
(**2**) R. Wildt, *Ap. J.*, **105**, 36, 1947.
(**3**) S. Chandrasekhar and G. Münch, *Ap. J.*, **104**, 446, 1946.
(**4**) S. Baumbach, *A.N.*, **263**, 121, 1937.

(5) W. Grotrian, *Z. Astrophys.*, **7**, 26, 1933.
(6) I. S. Shklovsky, *Russ. A. J.*, **25**, 145, 1948.
(7) H. C. van de Hulst, *Nature*, **163**, 24, 1949.
(8) G. Cillié and D. H. Menzel, *H.C.O. Circ.*, 410, 1935.
(9) W. H. McCrea, *M.N.*, **89**, 718, 1929.
(10) R. G. Giovanelli, *M.N.*, **109**, 372, 1949.
(11) S. Chapman and T. G. Cowling, *Mathematical Theory of Non-Uniform Gases*, p. 179, Cambridge, 1939.
(12) H. Alfvén, *Ark. Mat. Astr. Fys.*, **27**, No. 25A, 1941.
(13) D. F. Martyn, *Proc. Roy. Soc.* A, **193**, 44, 1948.
(14) E. Jahnke and F. Emde, *Tables of Functions with Formulae and Curves*, New York, 1943.
(15) M. Nicolet, *Mem. Inst. R. Met. Belg.*, **35**, 1949.
(16) W. N. Christiansen, D. E. Yabsley and B. Y. Mills, *Nature*, **164**, 569, 1949.
(17) J. L. Pawsey and D. E. Yabsley, *Aust. J. Sci. Res.*, **2**, 198, 1949.
(18) H. C. Minnett and N. Labrum, *Aust. J. Sci. Res.*, **3**, 60, 1950.
(19) J. H. Piddington and J. V. Hindman, *Aust. J. Sci. Res.*, **2**, 524, 1949.
(20) M. Laffineur and J. Houtgast, *Ann. d'Astrophys.*, **12**, 137, 1949.
(21) Serv. Prév. Ionosphérique Marine, *Obs. Solaires Av.-Juin* 1948, p. 14.
(22) J. H. Piddington, unpublished material.
(23) R. O. Redman, *M.N.*, **102**, 140, 1942.
(24) L. Goldberg, *Ap. J.*, **89**, 673, 1939. D. H. Menzel, *Pub. Lick. Obs.*, **17**, 1931.
(25) Reference (8) as modified by Wildt; Reference (2), p. 52.
(26) K. Wurm, *Zs. f. Astrophys.*, **25**, 109, 1948.
(27) R. G. Giovanelli, *Aust. J. Sci. Res.*, **1**, 275, 289, 305, 360, 1948.
(28) R. N. Thomas, *Ap. J.*, **109**, 480, 1949.
(29) C. R. Davidson, M. Minnaert, L. S. Ornstein and F. J. M. Stratton, *M.N.*, **88**, 536, 1928.
(30) S. E. A. van Dijke, *Ap. J.*, **99**, 121, 1944.
(31) J. Söderqvist, *Ark. Mat. Astr. Fys.*, **30**, No. 11A, 1944.
(32) D. R. Bates and H. S. W. Massey, *Proc. Roy. Soc.* A, **187**, 261, 1946.
(33) C. W. Allen, *Terr. Mag.*, **53**, 433, 1948.
(34) C. E. Moore, " Atomic Energy Levels," *Circ. N.B.S.*, 467, Washington, 1949.
(35) This was pointed out in correspondence with Mr D. R. Bates.
(36) D. R. Bates and A. Damgaard, *Phil. Trans. Roy. Soc.* A, **242**, 101, 1949.

PHOTOELECTRIC OBSERVATIONS OF S ANTLIAE*

A. R. Hogg and P. W. A. Bowe

(Communicated by the Commonwealth Astronomer)

(Received 1949 December 2)

Summary

Photoelectric observations have been made of the eclipsing binary S Antliae, using a blue-sensitive electron multiplier cell. A light curve has been derived showing the system to be one of two highly elliptical components. The system displays an anomalous reflection effect. A "uniform" solution has been obtained for the observations and the results compared with figures given by A. H. Joy. The present results differ in several respects from those obtained by Joy.

1. *Introductory.*—The desirability of observing S Antliae ($\alpha = 9^h 29^m.7$, $\delta = -28° 11'$ (1900)) by photoelectric methods has been recently pointed out by N. L. Pierce.† A summary of the earlier photometric observations has been given by A. H. Joy‡, who made spectrographic observations and demonstrated the binary nature of the object. Previously it had been held that the light variations were similar to those of a single rotating ellipsoid, having a period only half of that to be adopted for a binary object.§ Joy experienced difficulty in measuring the spectrum on account of a large rotational effect producing hazy lines. However, satisfactory velocity curves were obtained for the two components separately. The velocity curve was compared with a light curve constructed from the combined photometric observations of Pickering‖ and Wendell¶, using the epoch and period of Luizet**, and a complete orbit was calculated. However, Joy concluded that the stars would repay further accurate photometric observation.

2. *Observational.*—The observations were made with a multiplier photocell attached to the 9-inch Oddie refractor. The cell (RCA Type IP21) was supplied with a potential of 90 volts per stage from hearing-aid batteries, mounted on the telescope. A slide diaphragm containing holes of various sizes was arranged at the prime focus of the telescope. A movable reflecting prism was inserted into the beam behind the diaphragm to enable the star image to be centred in the appropriate hole with the aid of a finding eyepiece. When making the actual photometric measurement the prism was withdrawn and the light leaving the hole was collected by a Fabry lens which projected an unfocused image of the illuminated objective on to the cathode of the photocell. These arrangements are quite usual.

The output current of the photocell was passed through a 50-megohm resistor arranged in the grid circuit of a 954 tube operated with reduced filament, screen and anode voltages (4·1, 7·5 and 7·5 volts respectively). These values were

* The actual observations are not reproduced here but will be available on request from the Commonwealth Observatory.

† N. L. Pierce, *Contrib. Princeton Univ. Obs.*, **22**, 30, 1947.

‡ A. H. Joy, *Ap. J.*, **64**, 287, 1926.

§ H. Shapley, *A.N.*, **194**, 353, 1913.

‖ E. C. Pickering, *Harvard Ann.*, **46**, 124, 1903.

¶ O. C. Wendell, *Harvard Ann.*, **69**, 44, 1909.

** M. Luizet, *A.N.*, **165**, 291, 1904.

chosen so as to give a grid current low in comparison with the measured photocurrent. The arrangement was biassed so that the anode current with the unilluminated multiplier amounted to some 125 microamperes. This current was partially balanced by an opposing current and passed through a microammeter to give a half scale deflection of about 25 microamperes. When the multiplier was illuminated, the current in the microammeter circuit was decreased. This decrease was now compensated by applying a measured alteration of bias to the 50-megohm resistor. This measured alteration (expressed in millivolts) was taken as being proportional to the light intensity. Whilst the precise form of the relationship between light intensity and voltage alteration was not determined for this particular apparatus, careful measurements on similar equipment had shown it to be linear to within better than $0^{m}\cdot001$ over a range of $3^{m}\cdot0$.

The cell had been mounted in a vacuum tank and was provided with arrangements for refrigeration by dry ice. It was found that these precautions were not necessary to obtain stability in the dark current at the degree of sensitivity required. The cell was therefore operated under ordinary atmospheric conditions, and with the cell in the dark, only slow drifts of the microammeter needle, quite insufficient to affect normal readings, took place. In the illuminated condition the needle oscillated to an extent dependent on the state of the atmosphere. A capacity placed in parallel with the 50-megohm resistor gave the instrument a time constant of about 1 second. Each observation consisted usually of five readings of the light intensity of the variable interposed between two sets of the same number of readings of the comparison star. Corrections for the sky background were determined and applied when necessary. The usual reading for S Antliae was 700 mv., and was recorded to the nearest millivolt.

3. *Observations.*—Comparisons of S Antliae with ζ^{2} Antliae, totalling 408 in number, were obtained on 14 nights between 1949 January 15 and 1949 April 10. Control measures (59) were made of the difference $\zeta^{2}-\zeta^{1}$ Antliae on the same 14 nights. The nightly averages of the control magnitude differences all lay between $0^{m}\cdot320$ and $0^{m}\cdot300$ with the exception of a single result of $0^{m}\cdot274$ obtained on the first night of the programme. Corrections for atmospheric extinction were determined and applied as described elsewhere.* As the average deviation of the nightly measures, excluding the first night, amounted to but $0^{m}\cdot005$, the main comparison star ζ^{2} Antliae was accordingly treated as being of constant light. The observations were assembled on a provisional period of $0^{d}\cdot64834$ (Luizet's period doubled) and the means shown in Table I were obtained. The figures referred to as "semi-normals" in the table were obtained by averaging the results in groups of eight in order of increasing phase. The semi-normals were then folded about the epoch of principal minimum and meaned in nearby pairs so as to give normals of 16 observations. Pairs were chosen, so that in most cases each normal was made up of a semi-normal before and a semi-normal after the principal minimum. Both semi-normals and normals are shown in Tables I and II.

Period.—Estimates of the period were made by several methods, viz.:—

(*a*) from the interval between two primary minima observed in the present series;

(*b*) from the interval between two secondary minima;

(*c*) from the interval between Luizet's epoch and the epoch of the present observations.

* A. R. Hogg, *M.N.*, **106**, 294, 1946.

TABLE I

Semi-Normals—S Antliae

Phase	Intensity	Phase	Intensity	Phase	Intensity
d		d		d	
0·003	0·630	0·292	0·716	0·493	1·000
·019	·661	·303	·687	·498	·991
·043	·720	·317	·670	·505	·985
·057	·823	·327	·665	·511	·985
·073	·880	·335	·672	·525	·969
·093	·919	·341	·682	·541	·933
·125	·990	·359	·766	·550	·909
·143	1·002	·372	·773	·559	·897
·171	0·996	·383	·805	·569	·874
·184	·976	·376	·841	·576	·852
·207	·941	·407	·877	·583	·823
·218	·920	·420	·897	·593	·780
·225	·905	·432	·917	·607	·731
·237	·882	·441	·938	·619	·693
·247	·855	·453	·955	·629	·652
·255	·832	·473	·974	0·639	0·629
0·269	0·779	0·485	0·895	...	...

Each entry is the mean of eight observations arranged in increasing order of phase. The measurements are expressed in light intensities adjusted so that the maximum light of the curve is 1·000.

For methods (*a*) and (*b*), estimates of the epochs were made directly from a curve drawn through the observations on the particular night. In method (*c*) the epoch of the present observations was derived from the curve of normals which has been assembled, using Luizet's period. Points were chosen in pairs on the freehand curve drawn through the normals. The points in each pair had the same light intensity but one was on the ascending and the other on the descending branch of the primary minimum. The mean of the phases of each pair of points was found; the grand mean of 10 pairs measured at equal intervals of $0^m{\cdot}05$ gave the phase of the minimum, which added to the adopted epoch gave the epoch of

TABLE II

Normals—S Antliae—Folded means of 16 observations—equally spaced before and after primary minimum

Phase	(Obsd.)	(O–C)	Phase	(Obsd.)	(O–C)
d	m	m	d	m	m
0·006	0·502	−0·001	0·207	0·065	+0·011
·019	·458	0·000	·217	·093	+0·004
·042	·348	+0·018	·226	·113	+0·009
·056	·240	−0·002	·239	·139	−0·005
·091	·105	+0·003	·250	·179	−0·017
·102	·089	+0·017*	·267	·253	−0·010
·124	·022	−0·009	·288	·370	+0·016
·143	·006	−0·005	·305	·412	−0·004
·156	·005	−0·004†	·315	·433	−0·007
0·173	0·016	+0·002	0·321	0·4343	−0·003‡

Mean residual = $0^m{\cdot}007$

* 16 observations before principal minimum.
† 24 observations before principal minimum.
‡ 8 observations before principal minimum.

primary minimum 2432920·7715 ± 0·0003 (G.M.H.T.). The periods derived from these methods with assumptions as to the number of intervening cycles are shown in Table III, in which the standard error for a night minimum is 0ᵈ·004 as determined by a method similar to the above.

TABLE III

	Epoch—Julian Day G.M.H.T. 2400000·0+		Assumed cycles	Period
	i	ii	iii	(i–ii)/iii
(*a*) Primary minima	32962·139	32936·203	40	0ᵈ·64840
(*b*) Secondary minima	33013·025	32955·978	88	0ᵈ·64826
(*c*) Luizet's epoch	32920·7715	10741·5248	34,209	0ᵈ·64834528
			34,209·5	0ᵈ·64833580
			34,210	0ᵈ·64832642

The mean of the two results dependent solely on the present observations (*a* and *b*) gives a period of 0ᵈ·64833, which may be compared with Luizet's double period quoted by Joy, 0ᵈ·64833872. Examination of the three figures derived by method (*c*) suggests that the minima used in the Luizet and the present results do not refer to the same eclipse and that probably the time of occurrence of primary minimum (used in the sense of the present observations) will for some time to come be well represented by

$$2432920{\cdot}7715 + 0{\cdot}6483358\,T \quad \text{G.M.H.T.},$$

i. e. the result derived by assuming that 34,209·5 cycles intervened between the observations of Luizet and the present. It should be pointed out that the apparent depths of the two minima differ by less than 0ᵐ·07.

Light curve, rectification.—The appearance of the curve drawn through the normals of Table II indicated a high degree of ellipticity in the components. The inter-eclipse observations could be represented by the usual form of equation, viz.:—

$$L = A_0 + A_1 \cos\theta + A_2 \cos 2\theta. \qquad (1)$$

A least squares solution, using the normals between the phase angles $\theta = 45°$ and $\theta = 135°$ (which limits appear to be free of eclipses), gave

$$A_0 = +0{\cdot}878 \pm 0{\cdot}001 \text{ (s.d.)},$$
$$A_1 = +0{\cdot}021 \pm 0{\cdot}003,$$
$$A_2 = -0{\cdot}014 \pm 0{\cdot}003.$$

In the conventional representation of a close eclipsing binary system as consisting of two prolate spheroids of revolution with their major axes coincident with the line of centres, A_0 and A_2 are functions of the ellipticity and gravity darkening of the components, whilst A_1 is a function of the reflection effect. When θ is measured from the deeper minimum, both A_1 and A_2 are negative in sign in the conventional model. In the present case it is clear that the conventional model does not strictly apply. However, it has been found that if the method of rectification described by H. N. Russell * is applied, then the resultant rectified

* H. N. Russell, *Ap. J.*, **104**, 152, 1946

curve shows a sensibly constant maximum light and reasonably symmetrical eclipses. The curve rectified in this manner has been adopted as a basis for the solution. Owing to the empirical nature of the rectification it has been thought undesirable to apply the more recent refined methods of solution; also any further resort to least squares appears unreasonable.

Light curve, eclipses.—Following Russell's earlier methods* a freehand curve was drawn through the rectified normals of the primary minimum and the phases corresponding to definite fractions of the total light loss were obtained. These phases were used with trial values of the function $C/D=\chi(k\alpha_0{}^{\frac{1}{4}})$ and the coefficients $\omega_1(n)$ and $\omega_2(n)$ (Russell's equation (g)) to find what value of $\chi(k\alpha_0{}^{\frac{1}{4}})$ best suited the light-variation curve. Some six trial values of $\chi(k\alpha_0{}^{\frac{1}{4}})$ between 1·60 and 1·75 were used and the value of the constant D calculated. Sums of the squares of deviations of D from the mean value in each trial were computed and it was found that the minimum sum corresponded to $\chi(k\alpha_0{}^{\frac{1}{4}})=1{\cdot}665$. Next, using Russell's equation (j), pairs of values of k and α_0 (k = ratio of radii of larger and smaller stars, α_0 = maximum depth of eclipse) conformable with the observed rectified light losses at the two minima were calculated on the hypotheses that the primary minimum was (a) an occultation, (b) a transit. For each of the pairs the relevant values of $\chi(k\alpha_0{}^{\frac{1}{4}})$ was obtained from Russell's Table III and the values of k and α_0 conforming to the value of $\chi(k\alpha_0{}^{\frac{1}{4}})=1{\cdot}665$ (as deduced from the light curve) was determined by graphical interpolation (Fig. 1) as:—

	k	α_0
(a) Primary minimum an occultation	0·41	0·83
(b) Primary minimum a transit	0·47	0·98

These values led by the usual expressions to estimates of L_1/L_2 (ratio of light of large star to light of small star) as 3·3 for the occultation and 8·1 for the transit hypothesis. Joy (*loc. cit.*) states that "spectrograms indicate that one star is brighter than the other in the ratio of approximately 2:1". This makes the occultation hypothesis appear preferable, so it is adopted in the sequel. The value L_1/L_2 adopted is that derived from the present light curve and is not the spectrographic value used by Joy. The adopted values of k and α_0 lead to the computed results shown in Table II. The rectification for this computed curve has been effected with the constants A_0, A_1 and A_2 of equation (1), with the exception that the observed value of A_0 has been arbitrarily decreased from 0·878 to 0·875 to give a better distribution of positive and negative residuals. The mean residual amounts to $0^m{\cdot}007$.

Dimensions of orbit, masses, etc.—Continuing to make use of the methods set out by Russell (*loc. cit.*), the figures for the relative orbit shown in Table IV were obtained.

The present figures differ in various respects from those of Joy. Perhaps the most noticeable difference is that Joy's results indicate that the primary minimum is a transit and not, as considered here, an occultation. The radial velocity curve of course clearly shows that at the adopted phase of $\theta=0$ (i.e. the adopted primary minimum) the fainter star is in front. However, it is possible, as a considerable time elapsed between the original photometric and the spectroscopic observations, that the minimum adopted by Joy as the primary minimum might actually have been the secondary minimum of the light curve. Periods may be

* H. N. Russell, *Ap. J.*, **36**, 385, 1912.

calculated from a combination of Luizet's results with Joy's spectrographic curve, with the assumptions that the adopted epochs of the two sets of observations refer to (*a*) the same class of minimum or (*b*) to a different class of minimum, i. e. that in the interval between the two sets (11235^{d}·857), either an even or an odd number of the Luizet half-cycles have occurred. The two periods thus obtained

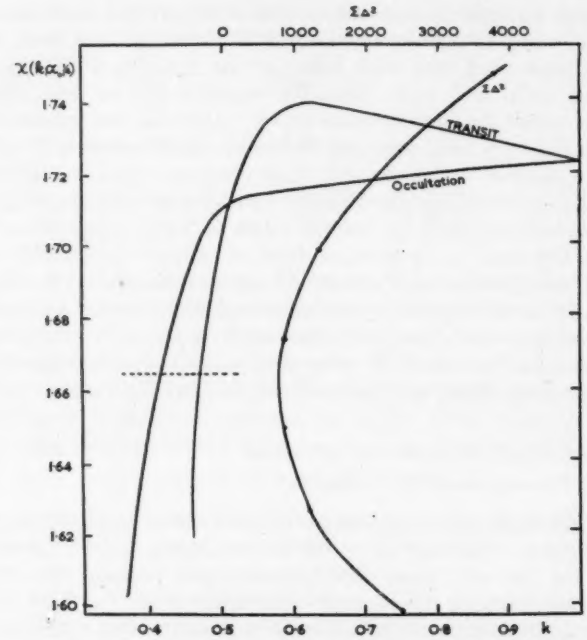


FIG. 1.—*The curve marked* $\Sigma\Delta^2$ *shows trial values of* $C/D=\chi(k\alpha_0\frac{1}{2})$ *in Russell's equation* (*g*), *plotted as ordinates against the corresponding values of the sum of the squares of the deviations of D from the mean obtained with each trial value of C/D. The minimum value of* $\Sigma\Delta^2$ *corresponds to the best trial value of* $\chi(k\alpha_0\frac{1}{2})$ *with respect to the curve for primary minimum.*

The possible values of $\chi(k\alpha_0\frac{1}{2})$ *demanded by Russell's equation* (*j*) (*relative depths of minima*) *on the hypotheses that the principal minimum is* (*a*) *a transit,* (*b*) *an occultation, are shown for various trial values of k in the curves marked accordingly. For reasons given in the text the occultation hypothesis is considered preferable and the intersection of the dotted line drawn for* $\Sigma\Delta^2$ *a minimum and the occultation curve gives the required value of k. The diagram places the uncertainty in k as about* 0·01 *units, which corresponds to an uncertainty* 0·005 *in* α_0.

are respectively (*a*) 0^{d}·64834720 and (*b*) 0^{d}·64832850. These figures differ from the Luizet period doubled by −8·5 and +10·2 millionths of a day. The point could best be cleared up by simultaneous or nearly simultaneous spectrographic and photometric observations. In the meanwhile it is reasonable to consider this to be the source of the discrepancy between the two periods, more especially since the difference in observed depths of the minima is only 0^{m}·07. The present results also differ from those of Joy in showing a greater degree of ellipticity, a smaller value of *k* and a more inclined orbit. The ratio of the surface brightness of the components (J_h/J_c) differs notably, but this is to be expected from the difference in the effective wave-lengths of the two sets of observations. The present observations were made with a cell having maximum response at 3900 A.

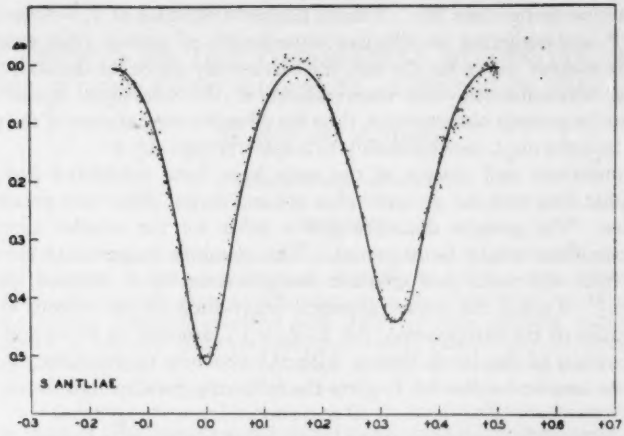


FIG. 2.—*The light curve of S Antliae. The dots represent individual observations and the curve is computed from the elements in Table IV.*

(makers' data), whilst the photometric observations used by Joy were taken in the visual range. However, the present results strongly suggest that the two components do not have the same spectral type. Joy states that the spectrum of both components is A8, although "there is much uncertainty in the case of the fainter spectrum". Taking the primary eclipse to be an occultation of a fainter and smaller star, then the present value of J_h/J_c suggests that this smaller star

TABLE IV

	Joy's results	Present results
	d	d
Period	0·64833872	0·6483358
Semi-axis major (large star) a_1	0·502	0·641
Semi-axis minor (large star) a_2	0·377	0·419
Semi-axis major (small star) b_1	0·389	0·289
Semi-axis minor (small star) b_2	0·292	0·172
Axial ratio $k=a_1/a_2=b_1/b_2$	0·775	0·410
Ellipticity coefficient z	0·340	0·506
Rectified light loss (primary min.)	0·145	0·194
Rectified light loss (secondary min.)	0·120	0·112
Ratio of surface brightness J_h/J_c	1·2	1·73
Light of large star L_1	0·67	0·77
Light of small star L_2	0·33	0·23
Inclination of orbit i (degrees)	62·2	70·1
Radius of relative orbit	$2{\cdot}30 \times 10^6$ km.	$2{\cdot}175 \times 10^6$ km.
Radius of primary orbit	$0{\cdot}817 \times 10^6$ km.	$0{\cdot}767 \times 10^6$ km.
Radius of secondary orbit	$1{\cdot}49 \times 10^6$ km.	$1{\cdot}40 \times 10^6$ km.
Semi-axis major (large star)	$1{\cdot}16 \times 10^6$ km.	$1{\cdot}39 \times 10^6$ km.
Semi-axis major (small star)	$0{\cdot}896 \times 10^6$ km.	$0{\cdot}630 \times 10^6$ km.
Mass (large star) $\odot=1$	0·75	0·63
Mass (small star) $\odot=1$	0·42	0·35
Density (large star) $\odot=1$	0·31	0·19
Density (small star) $\odot=1$	0·38	1·37
Absolute magnitude (large star)	2·9	2·8
Absolute magnitude (small star)	4·6	4·1
Parallax (seconds)	0·017	0·014

may be of a type earlier than A8. Taking Kuiper's estimate of $T_e = 8000$ deg. for an A8 star * and assigning an effective wave-length of 4000 A. (this estimate is based on the makers' curve for the cell, the black-body curve for the star, and the atmospheric transmission; the transmission of the telescopic optics is not included), to the present observations, then the effective temperature of the smaller star would be 9200 deg., corresponding to a spectral type A3. †

The dimensions and masses of the stars have been calculated from Joy's spectrographic data with the present value of i and do not differ very greatly from Joy's results. The present densities give a value for the smaller component rather greater than might be expected. The absolute magnitudes have been calculated from the radii and effective temperatures by a method given by Gaposchkin. ‡ Taking the visual apparent magnitude of the system as $6^m{\cdot}74$ the magnitudes of the components, for $L_1/L_2 = 3{\cdot}1$, amount to $7^m{\cdot}02$ and $8^m{\cdot}34$. The combination of the latter figures with the absolute magnitudes calculated with different assumed values of T_e gives the following parallax figures:—

T_e		π	
Large star	Small star	Large star	Small star
8000 deg.	8000 deg.	$0''{\cdot}013$	$0''{\cdot}017$
8000 deg.	9200 deg.	$0''{\cdot}013$	$0''{\cdot}015$

It will be seen that a readjustment of T_e, i.e. the spectral types of components, improves the agreement of the separately calculated values of the parallax. The mean of the values of the A8 + A3 system, $0''{\cdot}014$, has been adopted as the parallax of the system and the absolute magnitudes have been calculated accordingly. The parallax has been given by Adams § as $0''{\cdot}007$ from spectrographic results, and as $0''{\cdot}014$ by D. B. McLaughlin || from a recomputation of Joy's results.

4. *Conclusion.*—The results obtained here are in no sense definitive, nor can the uniform solution offered here be regarded as completely satisfactory, because of the anomalous reflection effect and the long interval between spectrographic and photometric results introducing uncertainty in the identification of spectrographic and photometric phases. However, the results are presented to draw attention to this most interesting eclipsing system with a high degree of ellipticity and to point out the desirability of carrying out a set of spectrographic observations before too great an interval has elapsed.

Commonwealth Observatory,
Mount Stromlo,
Canberra:
1949 *November* 23.

* G. P. Kuiper, *Ap. J.*, **88**, 428, 1938.

† A. H. Joy (private communication), however, considers that it is unlikely that the secondary is of an earlier type than the primary.

‡ S. Gaposchkin, "Variable Stars", *Harv. Obs. Monographs*, **5**, 50, 1938.

§ W. S. Adams, *Ap. J.*, **53**, 56, 1921.

|| D. B. McLaughlin, *Astron. Journ.*, **38**, 22, 1927.

THE EFFECTS OF THE TERRESTRIAL IONOSPHERE ON THE RADIO WAVES FROM DISCRETE SOURCES IN THE GALAXY

M. Ryle and A. Hewish

(Received 1950 April 24)

Summary

Observations of the discrete sources of radio waves in the galaxy have shown the existence of irregular refraction processes in the terrestrial ionosphere. These irregularities cause rapid fluctuations in the intensity of the radiation at the ground, whilst observations with aerials of high resolving power have shown, in addition, that the apparent position of a source may vary irregularly by 2–3 minutes of arc. The incidence of these irregularities shows a marked diurnal variation having a maximum at about $01^h\,00^m$ local time.

It does not seem possible to account for the irregularities in the ionosphere in terms of solar emissions, and an alternative mechanism is proposed which is based on the interception of interstellar matter moving under the gravitational attraction of the Sun. If this hypothesis is correct, further experiments may provide information of interest in theories of the accretion of matter by the Sun.

---

1. *Introduction.*—In a previous communication (Smith, 1950) an account was given of two series of experiments designed to determine the origin of the occasional fluctuations of intensity from the discrete sources of radio waves in the galaxy. One of these experiments (in which a wave-length of 3·7 m. was used to observe the source in Cygnus) was carried out in conjunction with the Jodrell Bank Experimental Station of Manchester University and some of the results have been described by Little and Lovell (1950). The second experiment was made with the radiation of wave-length 6·7 m. emitted from the major source in the constellation of Cassiopeia. Records of the intensity of the radiation were made simultaneously at two points, whose separation was varied.

The results of these experiments have been summarized by Smith (1950), who concluded that two types of fluctuations occur:—

(*a*) Sudden brief enhancements of the intensity which are suggestive of the "bursts" of solar radiation, and which appear to be due to genuine variations of the emission from the source.

(*b*) Irregular variations, probably due to refraction processes occurring in the outer regions of the terrestrial ionosphere.

In the present paper the latter phenomenon will be examined in greater detail, and an account will be given of further experimental results which have now been obtained.

The occurrence of fluctuations which produce variations of intensity both greater and less than the undisturbed value are suggestive of an irregular refraction mechanism. The irregular distribution of intensity which would be produced by such a mechanism is conveniently thought of in general terms as a diffraction pattern. Booker, Ratcliffe and Shinn (1950) have recently considered some ionospheric problems in this way, and some of their concepts will be used in this paper.

Diffraction of radio waves might occur either in the interstellar medium or in the terrestrial ionosphere, and it is first necessary to determine which of these two possible regions is responsible for the fluctuations in the intensity of the galactic sources. In the earlier experiments, which have been described by Smith (1950), simultaneous observations were made on a wave-length of 6·7 m. with two spaced receivers; it was found that the fluctuations, which had an average duration of 30 secs., were markedly different when the receivers were separated by a distance of 20 km., whilst no correlation between the two records could be detected when the separation was increased to 170 km. (in an east–west direction).

Similar results were obtained by Little and Lovell (1950) on a wave-length of 3·7 m.; they found small differences in the records when the separation of the receivers was only 3·9 km.

If the fluctuations were caused by diffraction in interstellar matter, the receivers would be moving through the diffraction pattern with the orbital velocity of the Earth, which is about 30 km./sec. The observation of fluctuations having a duration of 30 secs. would therefore imply that the diffraction pattern had a structure whose dimensions were at least 900 km. The lack of correlation between the records obtained with receivers at much smaller distances shows that the diffraction pattern is of considerably smaller dimensions; the observations can therefore only be explained if the diffraction pattern is moving with the Earth. It may therefore be concluded that the diffraction occurs in the Earth's atmosphere and is possibly due to irregularities in the density of ionization of the ionosphere.

It was also pointed out by Smith (1950) that there was no detailed correlation between the occurrence of the fluctuations and the occurrence of ionospheric or magnetic storm activity. During periods of great storm activity the fluctuations have frequently been observed to increase, but their amplitude was not greater than that occurring on some magnetically undisturbed days.

In the present paper the results of two separate investigations at Cambridge will be described. In one, observations were made on a wave-length of 3·7 m., using a system analogous to a transit telescope; with this instrument it is possible to make daily observations of the fluctuations of the intensity from a number of different sources, and continuous records of this type have been maintained for a period of 15 months.

The second series of observations was made on a wave-length of 6·7 m. and was intended to record the fluctuations of intensity from a single source continuously throughout the day. In these observations an aerial system of comparatively small aperture was directed at the pole, so that the intense source in Cassiopeia could be recorded continuously. These observations have been continued for a period of six months.

The two series of experiments will first be described in detail; the results obtained will then be discussed in relation to existing knowledge of the ionosphere, and finally possible mechanisms for the maintenance of irregular ionization will be considered.

2. *Observations on 3·7 m. with the "Transit" instrument.*—The aerial array used in these observations is arranged to receive uniformly over an aperture having a width of 20λ in an east–west direction, and a width of 1λ in a north–south direction. It is mounted horizontally and therefore receives in a narrow "fan" beam along the meridian; this "fan" beam has a width (to half sensitivity) of approximately $\pm 1\frac{1}{2}^\circ$ in an east–west plane, and extends from the pole to about

declination +10°. As the Earth rotates, the reception pattern will therefore sweep out the greater part of the northern hemisphere and radiation from any particular source will be received when it is situated within $\pm 1\frac{1}{2}°$ of the meridian; it is thus possible to observe each source for a period of about 20 minutes each day.

In some of the observations two similar aerials, separated by a distance of 110 wave-lengths, were used together to produce, by interference, a reception pattern of considerable resolving power (Ryle, 1950). The improved discrimination of this system makes it easier to study the fluctuations of the weaker sources against the general background radiation from the galaxy, although the system is less convenient for determining the detailed structure of the temporal variations of intensity. (The sensitivity of such a system varies periodically as the source moves across the reception pattern, and since the regular variations have a duration comparable with that of the irregular fluctuations it is difficult to deduce the precise shape of the irregular fluctuations from the record.)

If observations with this instrument are continued for a year, it becomes possible to find how the fluctuations from each source depend on local time, independently of any effects caused by change of the angle of incidence of the radiation, since each source is observed only near meridian transit. By making observations on a number of different sources at approximately the same declination but having different right ascensions, it becomes possible to separate effects associated with local time from those associated with annual changes.

If sources at different declinations are observed, it is further possible to estimate the effect of the angle of incidence of the radiation on the ionosphere.

Regular observations have been made with this instrument over the period 1948 December to 1950 March, and the results obtained will now be summarized. In analysing the records an "index of fluctuation" has been used to represent the degree of variability, which is defined as the ratio of the mean deviation of the intensity to the mean intensity. In this analysis no account has been taken of the variations of the average duration of the fluctuations observed on different occasions; the records have, however, shown that wide variations do occur, and that fluctuations having an average duration as short as three seconds and as long as five minutes are occasionally observed. This result will be considered later in Section 6.

Owing to the large day-to-day variations of the "index of fluctuation", the initial analysis is based on the mean value obtained for each month. The result of plotting this monthly mean for the period 1948 December to 1950 March, for four different sources, is shown in Fig. 1. The approximate positions of the four sources used in this analysis are:—

I. 19·01 (Cygnus), R.A. $19^h 57^m$; Decl. +40°.
II. 23·01 (Cassiopeia), R.A. $23^h 21^m$; Decl. +58°.
III. 05·01 (Taurus), R.A. $05^h 32^m$; Decl. +22°.
IV. 12·01 (Coma Berenices), R.A. $12^h 28^m$; Decl. +14°.

It can be seen that both the shape and the absolute magnitude of the curves obtained for the four different sources are remarkably similar, but that they are displaced in time in order of their respective right ascensions. It is now clear why the early results obtained by Ryle and Smith (1948) during the months of May and June 1948 had indicated that the source in Cassiopeia showed considerably smaller fluctuations than that in Cygnus.

The recent observations suggest that there may be a marked diurnal variation of the disturbing influence; an apparent annual variation would then be caused by the difference in local time at which observations were made.

In order to confirm this hypothesis, the four curves have been re-plotted in Fig. 2, with the time of observation as abscissae, allowance being made for the different right ascension of each source.

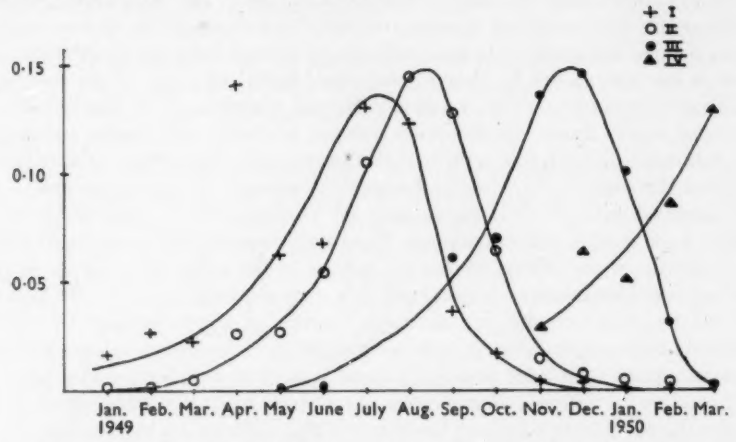


FIG. 1.—*Variation of the "fluctuation index" for four sources during the period* 1948 *December to* 1950 *March.*

Regular observations on source III (05·01) *were not started until* 1949 *May and observations of source IV* (12·01) *were not started until* 1949 *November.*

Ordinates: Fluctuation index.

It can be seen that the curves coincide within the limits of accuracy of the experiment. The results shown in Fig. 2 therefore suggest that there is a marked diurnal variation of the influence causing the fluctuations; this variation shows a rapid rise from $20^h\ 00^m$ to $22^h\ 00^m$, has a maximum at about $01^h\ 00^m$, and subsequently decays to a small value at noon.

Since the four sources are distributed widely in right ascension, it is clear that there can be no comparable annual variation of the disturbing influence; in particular it does not seem possible to account for the marked asymmetry of the curves in terms of an annual variation superimposed on a diurnal variation.

It is important to notice that at the times of observation the radiation from the different sources is incident on the ionosphere at angles between 6° and 38°, whereas the observed maximum values of the fluctuation index for the different sources are very similar; it may therefore be concluded that for angles less than 38° the angle of incidence does not have a major effect on the amplitude of the fluctuations. During these observations the larger angles of incidence were obtained for sources situated south of the zenith.

3. ***Continuous observations of the intense source in Cassiopeia on a wave-length of*** *6·7 m.*—In the second series of observations an aerial system of comparatively small aperture was directed at the pole so that the source in Cassiopeia (declination +58°) would always be within the reception pattern. For convenience in discriminating against the background radiation, two such aerials were used to

form an interference system of small resolving power. For certain positions of the source its angular motion will be perpendicular to the line joining the two aerials, and no periodic variation of the recorded power will be produced. Observations with this system therefore have poor discrimination against the general background radiation for two periods (of about two hours) each day. Since it would be difficult to make an accurate determination of the fluctuation index at these times, a second instrument was set up, having a similar pair of aerials but separated on a line perpendicular to that of the first pair. With these two instruments it becomes possible to make an accurate determination of the fluctuation index at all times of day.

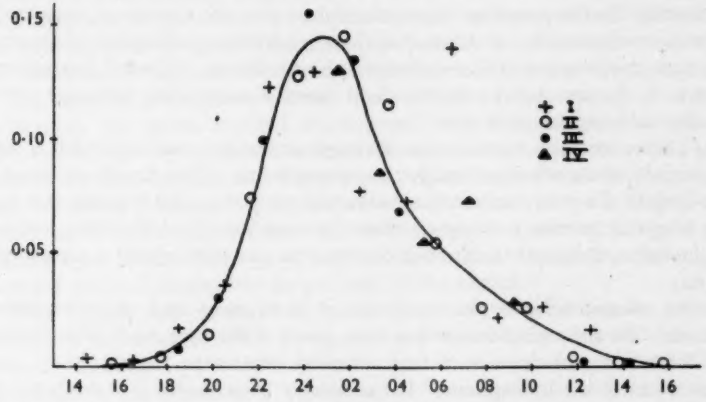


FIG. 2.—*Variation of the "fluctuation index" for the four sources plotted as a function of the time of observation.*

Ordinates: Fluctuation index.
Abscissae: Local time.

Observations with these instruments therefore make it possible to obtain a continuous record (apart from the disturbing effects of solar radiation and man-made interference) of the radiation from a single source. The variations in the fluctuation index throughout any one day are likely to be due to the combined effects of the diurnal variation and the variation of the angle of incidence of the radiation on the ionosphere. By combining the observations made over a period of six months, it becomes possible to determine the two effects separately.

Before carrying out observations at different angles of incidence, it is important to consider any possible effects due to refraction in the troposphere. Calculations based on the work of Appleton (1947) showed that in addition to any effects due to ionospheric refraction, observations made at angles of incidence greater than 80° might be appreciably affected by refraction in the troposphere. Since the superimposition of this additional process might confuse the study of the ionospheric refraction, it was decided to restrict the observations to the intense source in Cassiopeia, whose zenith angle is about 6° at upper culmination and 70° at lower culmination. This range of angles of incidence should be adequate to determine the dependence of the fluctuation index on the angle of incidence.

Observations of the source in Cassiopeia with this system have been maintained over the period 1949 October to 1950 March. The results are not yet

sufficiently complete to allow a full analysis of the relative importance of the diurnal effect and the angle of incidence in determining the amplitude of the fluctuations at a particular time, but it is clear that the main conclusions reached from the results on 3·7 m. are confirmed. Thus there appears to be a marked diurnal variation whose maximum occurs at about $01^h\ 00^m$ and there is some evidence for the same asymmetry. In addition, however, it is clear that the fluctuations are of greater amplitude when the source is at a large zenith angle, and observations made at the same local time near upper and lower culmination indicate that the fluctuations in the latter case are of approximately twice the amplitude of those in the former.

It should be noted that in these observations the large zenith angles occurred in the north. In the previous observations on 3·7 m. the sources at large zenith angles were to the south. It is possible that, in addition to the effect of the angle of incidence, the direction of observation may be significant—if, as will be suggested in Section 6, the terrestrial magnetic field exerts a controlling influence on the mechanism of irregular ionization.

The ability to observe one source throughout the day makes possible a more detailed study of the effects of magnetic storm activity. The results obtained on a wave-length of 6·7 m. confirm those obtained on 3·7 m., and indicate that even intense magnetic storms produce an effect not significantly greater than a typical midnight value, although the fluctuations may persist throughout a greater part of the day.

4. *The relation between the occurrence of fluctuations and other ionospheric phenomena.*—So far no indication has been given of the frequency of occurrence of the disturbed periods, nor of their detailed relation to the results of other measurements of the ionosphere. It has merely been shown that during a day when fluctuations are observed, the maximum effect occurs soon after midnight, with a gradual decrease throughout the early morning; there appears to be no detectable annual variation. It is now necessary to examine other evidence for the existence of irregular ionization in the ionosphere.

It is well known to ionosphere workers that pulse methods of examining Region F sometimes show remarkable irregularities in the reflecting region; this phenomenon, which is known as "spread-F", suggests that irregular regions of increased ionization density occur above the normal height of maximum ionization. It is possible that such clouds are frequently present, but that they are of insufficient density to be detected by pulse-reflection methods on any wave-length capable of penetrating the normal Region F.

Booker and Wells (1938) have examined this phenomenon, and have suggested that the experimental observations using pulses at relatively longer wave-lengths between 30 m. and 300 m. could be accounted for by Rayleigh scattering from spatial irregularities in the upper levels of Region F, in which the electron density was two or three times the undisturbed value. The occurrence of these irregular echoes shows a marked diurnal variation, usually setting in between $19^h\ 00^m$ and $20^h\ 00^m$ local time, and decreasing during the latter half of the night, with a rapid disappearance near dawn. Booker and Wells account for the latter phenomenon by supposing that the normal mechanism of Region F ionization, due to solar radiation, produces a continuous reflecting layer below the irregular region; any wave which could be scattered by the irregularities will then be reflected by this continuous layer and spread echoes will not be detectable. They also note

that whereas the phenomenon is very frequent, no obvious correlation with magnetic activity can be found.

Additional evidence for irregular ionization in Region F has been provided by radio pulse observations of the Moon on a wave-length of 14·8 m. (Kerr, Shain and Higgins, 1949) and by observations of the intense 5 m. radiation emitted by sunspots (Payne-Scott and McCready, 1948). In the latter case there was of course the possibility that the ionosphere was itself disturbed as a result of the enhanced solar activity.

The presence of irregularities of the type described by Booker and Wells might be sufficient to account for both of these observations and for the fluctuations in the radiation from discrete galactic sources. It is therefore important to examine the occurrence of spread-F echoes in relation to the occurrence of fluctuations in the intensity from the galactic sources; this possible relationship is also suggested by the similarity between the diurnal variations of the two phenomena.

An analysis has therefore been made of the period 1949 June 1 to 1949 September 30, when marked fluctuations were frequently observed in the intensity of the 3·7 m. radiation from the two intense sources I and II (in the constellations of Cygnus and Cassiopeia). Information concerning the occurrence of spread-F echoes during this period was obtained from the Radio Research Station at Slough. Marked fluctuations of the 3·7 m. radiation from the galactic sources were observed on 33 per cent of the nights during this period, whilst spread-F echoes occurred on 25 per cent of the nights.

The Slough data consisted of hourly observations indicating the presence or absence of spread-F echoes; since each source was only observed for a period of about 20 minutes at transit, an interval of three hours, centred at each of the transit times, was considered. A figure, in the range 0–3, was thus obtained for the spread-F activity, according to the number of occasions when spread-F was reported.

The observations therefore provide two independent series of simultaneous measurements of spread-F activity and fluctuation index. The correlation coefficient for each of these series may therefore be calculated, and the values obtained for sources I and II were 0·42 and 0·45. In each case the probability of such a figure occurring by chance (if the two phenomena were unrelated) is about 1 : 1000.

There can therefore be little doubt that both phenomena are related to the same primary disturbance and, as will be shown in the next section, there is good reason to believe that both phenomena could be accounted for by the same irregularities of ionization in Region F.

It is now necessary to examine what conditions in Region F would give rise to the observed fluctuations.

5. *The diffraction of plane waves by an irregular medium.*—In this section an attempt will be made to deduce from the experimental observations the nature of the irregularities in Region F. No attempt will be made in the present paper to carry out a rigorous analysis of the diffraction phenomena involved, and the aim will be merely to arrive at an order of magnitude for the dimensions and density of the irregular regions.

For wave-lengths in the range 3·7–6·7 m. absorption in the ionosphere is negligible and the effect of irregularities in the ionization density will be to produce irregular variations of phase across the emergent wave-front; the effect

of an irregular ionosphere may therefore be compared to the transmission of light through a transparent plate of glass of irregular thickness. In the following analysis the approximate width and depth of the irregularities will be deduced.

The problem of the diffraction of plane waves by an irregular screen has already been considered in some detail by Booker, Ratcliffe and Shinn (1950); their conclusions will be applied to the present problem.

For a diffracting screen which produces irregular variations of the intensity across the emergent wave-front, but which causes no variations of phase, it is possible to associate the spatial distribution of intensity across a remote plane parallel to the screen, with the irregular variations in the diffracting screen itself. It has already been pointed out that at these wave-lengths spatial irregularities in the ionization density will produce variations of phase across the emergent wave-front, but that the variations of intensity will be negligible. Under these conditions it is only possible to make deductions about the structure of the irregularities of ionization density by observing how both the intensity and the phase of the wave vary across the diffraction pattern at the ground.

Whilst it is relatively simple to deduce the instantaneous variation of intensity across the ground by making simultaneous observations of the intensity with two or more spaced receivers, the determination of the relative phase of the radiation incident at two widely separated points presents considerable difficulty.

In Section 5 (*a*) an account will be given of a method which has been used to determine the gradient of phase with distance along the ground, by instantaneous observations of the direction of the incident wave-front. In order to deduce the total variation of phase across the diffraction pattern at the ground, it is necessary to know the approximate dimensions of the region over which the measured gradient of phase is maintained. It is here necessary to consider the possible mechanisms which could give rise to the temporal variations of intensity (or of wave-front direction) at a point on the ground.

Two mechanisms suggest themselves:—

(i) The formation and decay of the additional ionization.

(ii) The movement of the diffraction pattern across the ground; this movement might be caused by the presence of winds in the ionosphere or simply by the rotation of the Earth.

From previous studies of the ionosphere it seems unlikely that the first of these mechanisms could give rise to fluctuations as rapid as those observed; the decay of ionization in Region F normally occupies a period of several hours. If the variations are caused by the movement of the diffraction pattern across the ground, a comparison of the average duration of the variations of intensity with that of the variations in the direction of the incident wave-front makes it possible to compare the average dimensions of the region over which the amplitude is constant with those of the region over which the direction of the wave-front is constant. From the previous observations it is then possible to determine the total variation of phase across the diffraction pattern at the ground; the methods of Booker, Ratcliffe and Shinn (1950) may then be used to determine the approximate structure of the irregularities of ionization density in the ionosphere.

(a) *The determination of the instantaneous direction of the incident wave-front.*—The instantaneous direction of the incident wave-front may be determined by observing the apparent position of the source; in the present series of experiments

a spaced-aerial interference system similar to that used by Ryle and Smith (1948) has been employed, and the variations in apparent right ascension have been observed.

In the initial experiments on 6·7 m. two aerials separated in an east–west direction by 10 wave-lengths were used to produce an interference pattern which was moved across the source in Cassiopeia by the Earth's rotation. In the absence of any refraction effects, the received intensity fell to zero each time there was a half-wave difference in the paths to the two aerials. If irregular refraction occurs in the ionosphere, it is to be expected that the direction of the wave-front will be modified; the times at which a minimum intensity occurs will therefore be displaced, and by observing the average displacement of a large number of the minima from their expected times, it is possible to determine the average variation in the direction of the wave-front at the aerial system.

In the initial experiments no measurable displacement of the minima could be detected even at times when marked fluctuations of the intensity were observed; it was therefore concluded that the direction of the wave-front was not undergoing variations greater than about 1°. Further experiments were therefore carried out, using a spacing between the aerials of 60 wave-lengths. During these observations occasional displacements were observed during very disturbed periods, which corresponded to an instantaneous deviation of the wave-front as great as $\frac{1}{2}°$; the average range of the observed deviations was 2–3 minutes of arc, as can be seen from Fig. 3 in which a histogram of the deviations is plotted for 50 observations.

It can therefore be concluded that no large deviations of the wave-front occur even when very large fluctuations of the intensity are observed; under normal conditions when fluctuations are present, the average deviation is 2–3 minutes of arc.

(b) *The total phase differences across the wave-front at the ground.*—The previous experiments indicated that the slope of the wave-front at a point on the ground was liable to vary through a range of a few minutes of arc. In these experiments observations were also made of the rate of change of the direction of the wave-front; it was found that the average rate of change was comparable with the rate of change of the intensity, i. e. the fluctuations had a duration of the order of 30 secs. On the assumption that the temporal variations of intensity at a point on the ground are due to the movement of the diffraction pattern across the receiver, it may therefore be concluded that the average dimensions of the regions over which the deviation of the wave-front is constant are comparable with those over which the intensity is constant. Previous experiments have shown this region to be about 5 km. in extent.

Knowing the average deviation of the wave-front (2–3 minutes of arc) and the average distance over which this deviation is maintained (5 km.), it may therefore be concluded that the total phase variation in the wave-front corresponds to a path difference of the order of 5 m., i. e. about one wave-length.

(c) *The dimensions of the irregularities in the ionosphere.*—If diffraction occurs at a non-absorbing "phase" screen (which merely modifies the phase distribution across the emergent wave-front), information concerning the structure of its irregularities may be deduced from observations of the distribution of phase and intensity across a remote plane which is parallel to the screen (Booker, Ratcliffe and Shinn, 1950). In the present problem it can be shown that the observations

on the ground are consistent with diffraction at an irregular ionosphere, in which the lateral dimensions of the irregularities are of the order of 5 km., and in which the average path difference introduced by the irregularities is of the order of one wave-length.

The total retardation introduced by Region F for a wave-length of 6·7 m. is of the order of 1000 wave-lengths; the observed fluctuations could therefore be accounted for by irregularities in the ionization density which correspond to a variation of the thickness of Region F of only 0·1 per cent.

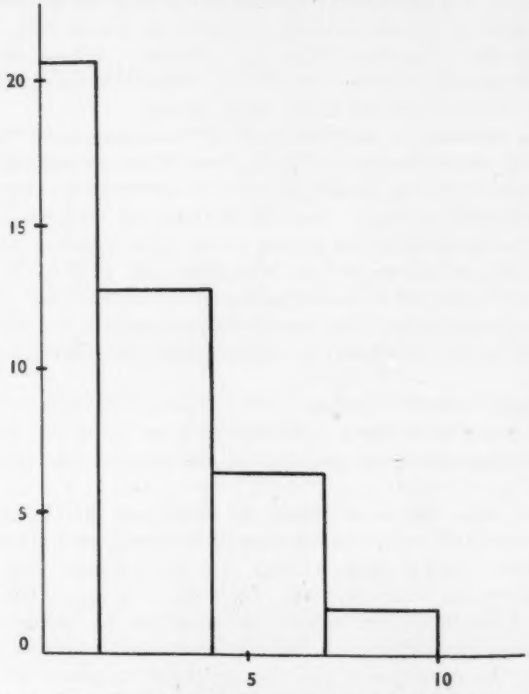


FIG. 3.—*Histogram of the instantaneous deviation of the incident wave-front (i. e. the error in the apparent R.A. of the source.)*
Ordinates: Number of observations.
Abscissae: Minutes of Arc.

In addition to the experiments carried out on a wave-length of 6·7 m., a limited number of observations were also made on 3·7 m. In these observations the average deviation of the wave-front was somewhat less, being of the order of 1–2 minutes of arc under normally disturbed conditions; taking account of the shorter wave-length, this result gives a variation of ionization density of the same order of magnitude as that found on 6·7 m.

6. *The origin of the irregular ionization.*—In the previous section it was shown that the observations could be accounted for by the presence of regions of enhanced ionization density in Region F, having average dimensions of the order of 5 km., and an additional ionization density corresponding to an increase in thickness of the region of about 0·1 per cent.

Previous studies of irregular ionization in Region F (Booker and Wells, 1938) have supposed that the diurnal variation in the occurrence of spread-F echoes was due to the formation at dawn of a continuous layer below the irregular region; this layer would then reflect back any radio pulses capable of being appreciably scattered by the irregular region, and would prevent further observation of the irregularities. The present series of experiments has been carried out on wavelengths sufficiently short to be unaffected by the formation of a continuous layer of ionization and therefore allows observation of the irregularities at all times of day.

These experiments have given the important result that there is a genuine diurnal variation in the occurrence of the irregular ionization; the irregularities do not exist during the afternoon, but build up rapidly between $20^h\ 00^m$ and $22^h\ 00^m$ to a maximum at about $01^h\ 00^m$. The time of maximum rate of growth appears to remain constant throughout the year, and does not, therefore, seem to be related to the time of sunset.

Two possible mechanisms for a diurnal variation of the irregular ionization suggest themselves:—

(*a*) A diurnal variation of some additional ionizing process, which has a maximum shortly after midnight.

(*b*) Diurnal changes in the ionosphere which alter the effectiveness of some permanent source of irregular ionization (i.e. the normal variations associated with solar radiation might at some periods of the day produce conditions particularly favourable for the formation of irregularities).

If the latter mechanism were responsible, however, it would be expected that the form of the diurnal variation would be similar to that of other ionospheric measurements, particularly of those relating to Region F. It is found that apart from the incidence of spread-F echoes, none of the normally recorded variables of Region F shows a similar diurnal variation. It therefore seems more likely that the diurnal variation of the fluctuation index is caused by actual variations in the intensity of the process causing irregular ionization.

The origin of the irregular ionization observed in Region E has been extensively studied (Appleton and Naismith, 1947; Lovell, 1948); the results have shown a conclusive relationship with the incidence of meteors. A similar origin does not appear possible for the irregularities in Region F, both because of the absence of any correlation with the major meteor showers, and because of the marked and regular maximum soon after midnight.

In addition to corpuscular matter of meteoric origin, it is well known that at times of solar activity intense streams of particles are incident on the Earth's atmosphere and give rise to magnetic storms and aurorae. It has been found that at such times additional fluctuations are observed which have a character similar to those normally observed, but which persist throughout a greater part of the day. Whilst this observation suggests that corpuscular matter may also be responsible for the normal fluctuations, it does not seem likely that such matter could be of solar origin.

It is possible that the incidence of corpuscular matter of interstellar origin might provide an explanation of the irregular ionization of Region F. It has been shown by Hoyle and Lyttleton (1939) that the gravitational attraction of the

Sun will cause interstellar matter to fall towards the Sun from a considerable distance (which depends on the velocity of the solar system relative to the interstellar matter). When the particles reach the orbit of the Earth they will have attained a radial velocity given by $\frac{1}{2}mv^2 = GM_{\odot}m/r$, or $v = \sqrt{(2GM_{\odot}/r)}$, where G is the gravitational constant, $M_{\odot}$ is the mass of the Sun, and r is the orbital radius of the Earth; the density of the particles may be considerably greater than that of the undisturbed interstellar matter.

Since the orbital velocity of the Earth is given by $\sqrt{(GM_{\odot}/r)}$, the relative motion of the interstellar matter will be normal to the surface of the Earth at about $02^h\ 20^m$ local time. (This result is only obtained if the general motion of the solar system relative to the interstellar matter is small compared with the orbital velocity of the Earth.) If interstellar particles were responsible for producing ionization in Region F, it might therefore be expected that the maximum disturbance would occur at about $02^h\ 20^m$ local time, with a cut-off somewhat before $20^h\ 20^m$ and somewhat after $08^h\ 20^m$. Although the time of the maximum disturbance does not agree with the experimental observations of the maximum fluctuation index (Fig. 2), the times of onset and disappearance of the phenomenon are in good agreement with those predicted by such a mechanism.

As has already been pointed out, it does not seem possible to account for the asymmetry of the experimental curve (Fig. 2) in terms of an annual effect, since all four sources exhibit the same asymmetry. It does not therefore seem possible to explain it in terms of the general motion of the solar system relative to the interstellar medium. It is suggested that the asymmetry might be caused by some local controlling effect, such as the deflection of the incoming particles by the terrestrial magnetic field.

Additional evidence for a mechanism based on the incidence of corpuscular matter is provided by the occasional observation of a sudden increase in both the amplitude and rate of the fluctuations; during a subsequent period of 30–40 minutes it is then found that the fluctuations gradually decrease in amplitude and increase in duration. This observation suggests the formation of localized intense regions of additional ionization, which gradually diffuse out to form a more uniform region.

In addition to the evidence provided by radio measurements, observations of the intensity of the light from the night sky (Vegard, 1948; Vassy and Vassy, 1948) have shown that marked irregularities are frequently observed; these irregularities are of brief duration and may cover only a small fraction of the sky and are usually observed at about $22^h\ 00^m$ and in directions to the north. These authors have noted the similarity of the irregularities to those associated with auroral activity, although they are far less intense, and are observed much more frequently. It seems possible that their observations may also be interpreted in terms of the incidence of corpuscular matter not necessarily of solar origin.

7. *Conclusions.*—Observations of the discrete sources of radio waves in the galaxy on wave-lengths of 3·7 m. and 6·7 m. have indicated that fluctuations in the received intensity are caused by irregularities in Region F of the terrestrial ionosphere. The occurrence of these irregularities shows a marked diurnal variation, the fluctuations increasing rapidly at about $21^h\ 00^m$ local time, to a maximum at $01^h\ 00^m$; the onset does not appear to be related to the time of sunset.

Experiments made with aerial systems of large resolving power have enabled the instantaneous direction of arrival of the radiation to be measured. It has been found that in addition to variations of the intensity, the apparent position of a source may show rapid variations; the average displacement of the position at times of marked disturbance may be as much as 3 minutes of arc at a wave-length of 6·7 m. and 1–2 minutes of arc at 3·7 m. These variations, which have an average duration of about 30 secs., are therefore sufficient to cause difficulties in the accurate location of the galactic sources at the longer wave-lengths. Observations with the present accuracy are not, however, likely to be affected during the daytime on these wave-lengths, and observations on shorter wave-lengths should be little affected at any time of day.

An analysis of the distribution of phase and intensity of the radiation across the ground makes it possible to obtain information on the irregular structure in the ionosphere which is responsible for the fluctuations; a detailed analysis of this problem will be given elsewhere, but it can be concluded from the present experimental results that the lateral dimensions of the irregularities are of the order of 5 km., while the variations of ionization density correspond to variations in thickness of Region F of about 0·1 per cent. These observations suggest an important extension to the present pulse methods of ionospheric study. In particular it becomes possible to obtain information on irregular regions situated above a uniform layer of greater ionization density; earlier pulse observations of irregular ionization in Region F were not able to demonstrate that the irregularities exhibited a marked diurnal variation.

The origin of the irregular ionization is not known; it does not, however, seem likely that it is due to solar emissions, owing to the form of the diurnal variation. One possible mechanism which provides a satisfactory explanation of the diurnal variation is based on the interception of interstellar matter which is moving in the gravitational field of the Sun. If this hypothesis is correct, the absence of any significant differences in the shape of the diurnal variation at different times of year would indicate that the general motion of the solar system relative to the interstellar matter is small compared with the orbital velocity of the Earth.

If this hypothesis is confirmed by more extended observations, the results may be of considerable interest in their application to theories of the accretion of matter by the Sun; it is possible that more accurate observations might then enable any general motion of the solar system relative to the interstellar medium to be determined.

Acknowledgments

This work forms part of a programme of research which is supported at the Cavendish Laboratory by the Department of Scientific and Industrial Research. One of us (A. H.) also holds a Maintenance Grant from the same source.

We would like to express our thanks to the Director of Radio Research (D.S.I.R.) for allowing us to use ionospheric data collected at the Radio Research Station, Slough, and to Mr Naismith for his assistance in providing and interpreting it.

Cavendish Laboratory,
Cambridge :
1950 *April* 19.

References

Appleton, E. V. (1947), *Phys. and Roy. Met. Soc. Report*, p. 1.
Appleton, E. V. and Naismith, R. (1947), *Proc. Phys. Soc.*, **59**, 461.
Booker, H. G., Ratcliffe, J. A. and Shinn, D. H. (1950), *Phil. Trans. Roy. Soc.*, **242A**, 579.
Booker, H. G. and Wells, H. W. (1938), *Terr. Mag.*, **43**, 249.
Hoyle, F. and Lyttleton, R. A. (1939), *Proc. Camb. Phil. Soc.*, **35**, 405.
Kerr, J. F., Shain, C. A. and Higgins, C. S. (1949), *Nature*, **163**, 310.
Little, C. G. and Lovell, A. C. B. (1950), *Nature*, **165**, 423.
Lovell, A. C. B. (1948), *Rep. Prog. Phys.*, **12**, 415.
Payne-Scott, R. and McCready, L. L. (1948), *Terr. Mag.*, **53**, 429.
Ryle, M. (1950), *Rep. Prog. Phys.*, **13**, 184.
Ryle, M. and Smith, F. G. (1948), *Nature*, **162**, 462.
Smith, F. G. (1950), *Nature*, **165**, 422.
Vassy, A. and Vassy, E. (1948), *Report of the Gassiot Committee of the Royal Society*, p. 60.
Vegard, L. (1948), *International Council of Scientific Unions (Report of Mixed Commission on Ionosphere)*, p. 111.

A STUDY OF THE SPECTRUM AND MAGNETIC VARIABLE STAR HD 125248

D. W. N. Stibbs

(Communicated by the Commonwealth Astronomer)

(Received 1950 March 7)

Summary

The relationship between variability in line intensity, radial velocity, magnetic field, and luminosity of HD 125248 is examined. Using the period 9·295 days given by Deutsch for the spectrum changes, the radial velocity measurements made by Morgan in 1931 are analysed. These measurements suggest a radial velocity variation of approximate amplitude 13 km./sec. with zero velocity when the *Eu* II lines have their maximum intensity. Thus the velocity zeros occur when the magnetic field variation discovered by Babcock has its extreme values. Photometric measurements leading to the light curve of the star are described. An harmonic analysis shows the light variation to be symmetrical about the maximum, the first and second harmonics being the principal terms. The elements of the light variation are

$$\text{Light maximum} = \text{JD } 2433\ 103{\cdot}95 + 9{\cdot}295E,$$

the maximum occurring when the *Cr* II lines have their greatest intensity. The light maximum occurs when the magnetic intensity of the polar field is −6500 gauss, and the minimum when the field is +7800 gauss. The amplitude of the light variation at the effective wave-length 4000 A. is consistent with a colour temperature change from maximum to minimum light of about 250 deg. K., corresponding to a gradient change of about 0·01. No change in colour has so far been detected.

---

1. The spectrum of HD 125248 (BD −18° 3789 Vir) has been analysed by Morgan (1) who discovered large changes in the absorption line intensities. The lines due to ionized europium were found to exhibit periodic variations in intensity out of phase with somewhat similar variations in the lines due to ionized chromium. The neutral chromium line λ 4254 was found to vary with the *Cr* II lines with a range of variation about the same as the strongest *Cr* II lines λλ 4242, 4269 and 4558. The most marked *Eu* II intensity variation was shown by λ 4205. An unidentified line at λ 4296 was found to vary in phase with the *Eu* II lines. Possible changes in the lines of *H* and *Mg* II were also noted.

The line intensity variations were later studied in detail by Deutsch (2) who found the elements of the variation to be

$$Eu\text{ II maximum} = Cr\text{ II minimum} = \text{JD } 2430\ 143{\cdot}07 + 9{\cdot}295E.$$

In addition to the above-mentioned changes in the spectrum the lines of *Ca* II K and *Mg* II λ 4481 were found to vary with *Cr* II, and lines of *Si* II and *Fe* II probably with *Cr* II also. Deutsch found an abrupt change in the spectrum near the phases $0^{P}{\cdot}2$ and $0^{P}{\cdot}8$ after *Eu* II maximum, the minimum being flat while the maximum is more rounded. He concludes that the line intensity variation is certainly anharmonic.

Somewhat similar spectrum variability is shown by α^2 Canum Venaticorum (HD 112413) and has been studied comprehensively by Struve and Swings (**3**). This star shows the same out-of-phase variation between the *Eu* II and *Cr* II lines as in HD 125248, but the range of variability with a period of 5·469 days is much greater, the *Eu* II lines increasing from invisibility to an intensity which places them amongst the most conspicuous lines in the whole spectrum.

2. Morgan found that the line intensity estimates from twelve plates were satisfied by a period of 3·18 days, but the radial velocity measures assembled on this period did not give any certain evidence of periodicity. However, the scatter in the measures appeared to be much larger than it should be if the velocity were constant. We have found that when the phases of Morgan's radial velocity measurements are calculated with the revised period given by Deutsch, there is then definite evidence of variability. Weighting the measures in his 1931 series, according to plate quality, we have made a least squares solution for the radial velocity in the form

$$V = v + a\sin\theta + b\cos\theta,$$

where $\theta = 2\pi \times$ phase, and phase zero = *Eu* II maximum. Twelve equations of condition gave

$$v = -2{\cdot}4 \pm 1{\cdot}2 \text{ (p.e.)}, \qquad a = 12{\cdot}6 \pm 2{\cdot}0, \qquad b = 1{\cdot}6 \pm 1{\cdot}4.$$

We conclude that there is evidence for a sine variation of amplitude about 13 km./sec., the variable component of the velocity being zero when the *Eu* II lines are a maximum.

It is not clear to which lines Morgan's radial velocity measures refer, and consequently the variation suggested is to be regarded as tentative in view of the fact that the velocity measurements of Struve and Swings (**3**) for α^2 C Ven show a complex state of affairs. The rare earths and some other elements of low ionization potential give a shallow minimum and a sharp maximum with zero velocity roughly in phase with the *Eu* II line intensity maximum; *Cr* II and some other elements give two maxima in the radial velocity per period, whilst *Mg* II, *Si* II, *H* and *Ca* II show no appreciable variation. Babcock (**4**) has indicated that in HD 125248 there are systematic changes in radial velocity with phase for three groups of elements characterized by the rare earths, iron and chromium, but that the differences between the various elements are small. For the purpose of the discussion in Section 5 we shall make use of our analysis of Morgan's measures which suggest an amplitude of about 13 km./sec. with zero velocity at the *Eu* II maximum.

3. The investigations of Babcock (**5**) on magnetic effects in stellar spectra have made it possible to measure the intensity of the polar field from a study of the Zeeman effect in the spectrum lines. The displacement δ of the centre of gravity of a normal Zeeman triplet blend when observed with a circular analyser is interpreted in terms of the surface field due to a magnetic dipole at the centre of the star. The Zeeman effect integrated over the visible hemisphere has been calculated by Babcock for the case of a star with limb darkening coefficient $u = 0{\cdot}45$, the magnetic axis and the line of sight being coincident. The integrated displacement, obtained partly by numerical methods, is given as $\delta = 0{\cdot}311 aH_p$, where aH_p is the displacement at the pole. The integrated Zeeman effect for a star viewed at any angle to the magnetic axis has not yet been given. In view of the possibility that the magnetic axis may be inclined to the axis of rotation, as in

the Earth and the Sun, we have given in an appendix to this paper a brief consideration of the case of a star with a linear law of darkening $(1-u+u\cos\theta)$ viewed at any angle α to the magnetic axis. The integrated displacement is found to be given by

$$\delta(\alpha, u) = a\frac{M\cos\alpha}{10R^3}\cdot\frac{15+u}{3-u},$$

where R is the radius of the star and M the moment of the magnetic dipole at the centre of the star giving rise to the polar field $H_p = 2M/R^3$. We conclude that a star whose surface field is due to a dipole of magnetic moment M at its centre, when viewed at an angle α to the magnetic axis, gives an integrated Zeeman effect due to a dipole of moment $M\cos\alpha$.

If the axis of rotation of the star makes an angle ρ with the line of sight and an angle β with the magnetic axis the value of α at any phase of the rotation is given simply by

$$\cos\alpha = \cos\rho\cos\beta + \sin\rho\sin\beta\cos\theta,$$

where $\theta = 2\pi \times$ phase. The variation in the displacement has been computed for a variety of conditions in order to demonstrate the order of magnitude of the magnetic variability which the movement of the magnetic axis simulates.

Babcock (**6**) has found that the magnetic field of HD 125248 is not only variable but reverses its polarity. The magnetic variation with period 9·3 days takes place in synchronism with the spectrum variations described in Section 1, showing a maximum polar field strength +7800 gauss in phase with the maximum in the *Eu* II line intensity, and a polar field −6500 gauss when the *Cr* II lines are strong. Zeros in the field occur approximately when the *Eu* II and *Cr* II lines are about the same intensity. Although the effect described in the preceding paragraph and illustrated in Fig. 1 may well be of importance in some cases of magnetic variability, as has already been suggested by Babcock (**4**), the reversal in polarity in HD 125248 and the fact that the spectrum lines show only about half the rotational broadening due to a nine-day rotation, suggest that the magnetic changes in this case are intrinsic variations of the surface field.

4. A photometric study of HD 125248 has been made at the Commonwealth Observatory with a view to detecting changes in the luminosity of the star. The investigation was undertaken following Dr Babcock's discovery of magnetic variability, a phenomenon which seemed to cast some doubt on the proposed relationship between angular momentum and magnetic moment. A preliminary announcement has been made by the writer (**7**) of the discovery of a small but well-defined light variation of range $0^m{\cdot}053$ with a period of 9·3 days. The observations were made with the Oddie Refractor (nine inch), and a 1P21 electron multiplier photometer which has been already described (**8**).

A survey of possible comparison stars for HD 125248 indicated the suitability of HD 124683, a star which combined the convenience of proximity with the virtue of constancy, and in addition was only one subclass different in spectral type from that of the variable. Table I gives the results of the photometric observations. Each comparison made consisted of five readings on HD 125248 taken between two sets of five readings on the comparison star. The magnitude differences HD 125248 minus HD 124683 for the night were averaged and the probable error determined, the mean epoch of the comparisons for each night being reduced to heliocentric Julian Day. A glance at the second and fifth columns in Table I

shows the continuity of the observing. The light variation was followed through eight periods in the twenty-six observing nights finally adopted for the determination of the light curve.

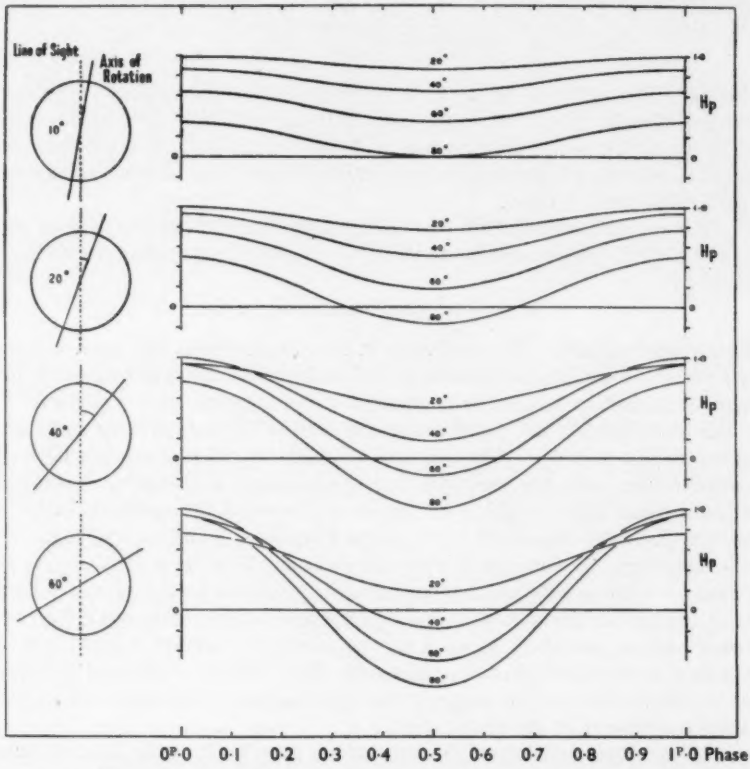


FIG. 1.—*Apparent magnetic variations during the rotation of a star when viewed obliquely to the axis of rotation, the magnetic and rotational axes not being coincident. The four curves shown for each inclination to the line of sight are for co-latitudes 20°, 40°, 60° and 80° of the magnetic pole.*

The well-defined nature of the variation shown by the preliminary light curve (7) strongly supports the nine-day period as the period of light variability. However, Deutsch has drawn attention to the possibility that a period of the order of the reciprocal of his adopted period is not entirely ruled out by the spectrum measurements, although the use of periods near one day does give an appreciably greater scatter in his assembled observations. But the photometric measurements enable us to reject this possibility because a period of order one day would give a maximum rate of change of brightness of about $0^{m}\cdot01$ per hour, and throughout an observing period extending over two to three hours at the epoch of maximum rate the residuals should readily show an effect. An examination of the behaviour of the residuals throughout each night's observations shows that in no instance do they have a definite trend and furthermore the expected effect is between ten and twenty times the probable error in the mean magnitude

TABLE I

Photometric Observations

Number of comparisons	Mean epoch of comparison JD 2433100·0+	Mean magnitude difference HD 125248 −HD 124683	Number of comparisons	Mean epoch of comparison JD 2433100·0+	Mean magnitude difference HD 125248 −HD 124683
17	6·928	0·355±002$_0$	18	40·910	0·314±000$_8$
15	7·933	·367±001$_0$	22	41·898	·318±000$_8$
14	8·950	·368±002$_2$	18	42·929	·339±001$_0$
11	9·994	·363±001$_6$	19	44·913	·362±001$_1$
10	18·991	·364±000$_8$	30	45·900	·365±000$_8$
17	23·927	·330±000$_8$	13	47·906	·352±001$_0$
9	26·927	·369±000$_8$	19	49·881	·321±000$_9$
12	28·893	·359±000$_6$	18	50·883	·317±000$_6$
10	33·894	·345±001$_8$	9	53·923	·364±002$_2$
17	34·895	·358±001$_1$	15	58·890	·320±000$_8$
21	35·916	·366±001$_0$	15	59·898	·314±000$_9$
20	37·927	·359±000$_8$	10	75·887	·353±002$_0$
12	39·885	0·325±000$_8$	14	76·884	0·332±002$_0$

difference for the night. The twenty-six mean points have therefore been combined into a composite curve on the basis of the period of 9·295 days given by Deutsch for the spectrum variations. The resulting light curve is shown in Fig. 2, the epoch of maximum light being JD 2433 103·95 with an estimated probable error of ±0·15 days. Table II gives the points on the light curve in order of increasing phase, zero phase being at light maximum. The magnitude difference Δm measured from maximum light is converted into light intensity on the basis of maximum light equal to unity.

TABLE II

Light Curve

Phase	Light intensity	Δm	O–C	Phase	Light intensity	Δm	O–C
P		m	m	m		m	m
0·019	1·0000	0·000	0·000	0·513	0·9541	0·051	+0·002
·049	·9972	·003	− ·001	·538	·9514	·054	− ·001
·083	·9963	·004	+ ·002	·618	·9550	·050	− ·001
·149	·9826	·016	+ ·001	·650	·9558	·049	− ·002
·194	·9772	·025	+ ·001	·655	·9594	·045	+ ·001
·222	·9718	·031	− ·001	·684	·9594	·045	− ·001
·320	·9629	·041	+ ·003	·729	·9655	·038	·000
·329	·9603	·044	+ ·001	·739	·9646	·039	− ·003
·376	·9550	·050	− ·001	·847	·9835	·018	·000
·407	·9568	·048	+ ·003	·866	·9900	·011	+ ·003
·428	·9523	·053	− ·001	·911	·9945	·006	+ ·001
·439	·9532	·052	·000	·942	·9936	·007	− ·004
0·472	0·9506	0·055	−0·002	0·976	1·0000	0·000	0·000

A preliminary harmonic analysis of the light variation showed that the coefficients of the sine terms were negligible, only the first three harmonics appearing in the cosine series. Accordingly a least squares solution was made for the coefficients in the equation

$$l = a + b\cos\theta + c\cos 2\theta + d\cos 3\theta,$$

where $\theta = 2\pi \times$ phase. Taking the observations to be of equal weight we find

$$a = 0{\cdot}9722 \pm 0{\cdot}0002 \text{ (p.e.)}, \qquad b = 0{\cdot}0234 \pm 0{\cdot}0003,$$
$$c = 0{\cdot}0039 \pm 0{\cdot}0003, \qquad d = 0{\cdot}0005 \pm 0{\cdot}0003.$$

The mean light per cycle is of course the constant term in the cosine series. The curve drawn amongst the observations is the light variation given by the harmonic analysis. The residuals O–C in Table II are given in magnitude difference.

The elements of the light variation are

$$\text{Light maximum} = \text{JD } 2433103{\cdot}95 + 9{\cdot}295\,E,$$

the light minimum being 0·9522, and the corresponding magnitude difference from the maximum being $0^{m}{\cdot}053$ in an effective wave-length of 4000 A. approximately.

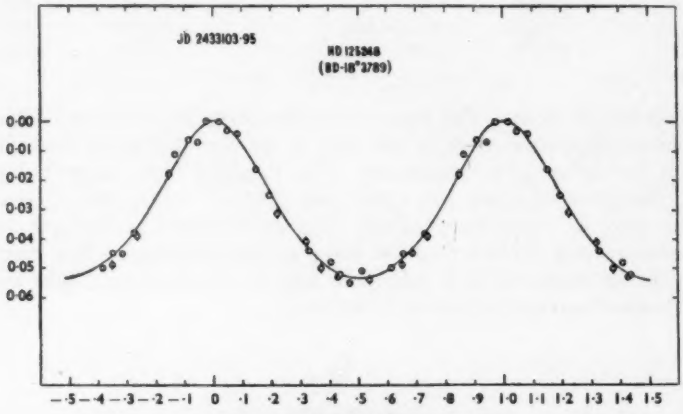


FIG. 2.—*HD 125248 light variability.*
Probable error of $0^{m}{\cdot}001$ is indicated by the circles. Mean points for which the p.e. exceeded this amount are indicated.
Ordinates: Magnitude difference.
Abscissae: Phase in period units.

5. We now discuss the relationship between the variation in line intensity, radial velocity, magnetic field and luminosity outlined in the preceding sections. From the elements of Deutsch for the variation in line intensity given in Section 1, and with $E = 318$ periods we find

$$Eu\text{ II maximum} = Cr\text{ II minimum} = \text{JD } 2433098{\cdot}88$$

for the extrapolated epoch of zero phase. Comparing this epoch with that of the light maximum in Section 4 we find that the phase of the light maximum with respect to the *Eu* II variation is $0^{P}{\cdot}55$. Thus the moderately sharp light maximum occurs approximately when the *Cr* II lines have their maximum intensity, and the somewhat shallow light minimum occurs at the *Eu* II maximum. Fig. 3 gives the relationship between the several types of variation. The spectral

characteristic given in the figure is that derived by Deutsch from Morgan's 1931 intensity estimates; the radial velocity variation refers to the same epoch but the phase of the light variation has been derived as above, by extrapolating Deutsch's period to the present time. To gain some idea as to the accuracy of the period $9^d{\cdot}295$ we notice that the epoch of maximum in the spectral characteristic which Deutsch derived from Morgan's estimates is located at JD 2426367·90, whereas the later observations in 1944/5 by Deutsch give the epoch of *Eu* II maximum at JD 2430143·07. If 406 periods have elapsed between these two values of epoch we obtain $P = 9^d{\cdot}298$. Now an uncertainty in the period as great as $\pm 0^d{\cdot}003$ would give an error of $\pm 0^d{\cdot}95$ in the present epoch of *Eu* II maximum obtained by extrapolation from JD 2430143·07 to the present time by means of 318 periods. Hence the uncertainty in the location of the phase of the light maximum with respect to the *Eu* II variations would be $\pm 0^P{\cdot}10$.

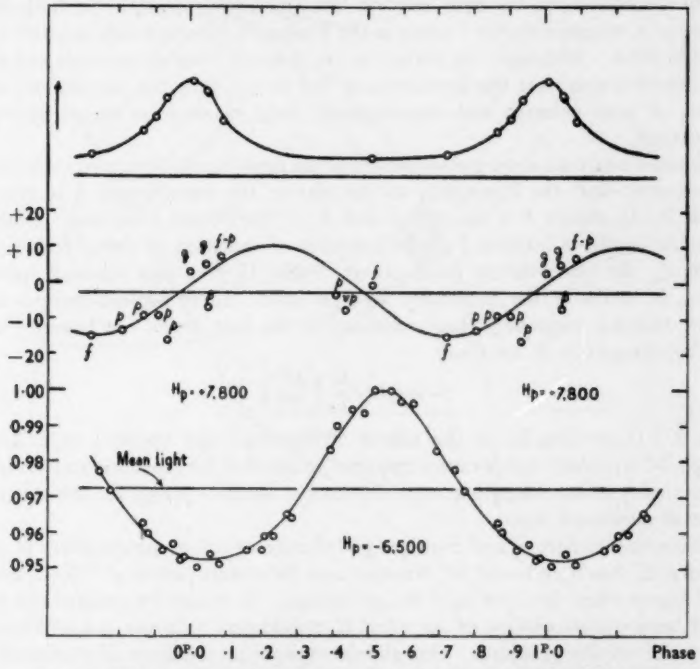


FIG. 3.—*HD* 125248. *Variability in line intensity, magnetic field and luminosity.*
Upper curve: Special characteristic. Eu II strong.
Middle curve: Radial velocity (km./sec.).
Lower curve: Light intensity.

Comparing the light variation with the radial velocity measures discussed in Section 2 we see that the light maxima and minima occur at phases of zero velocity. The magnetic variations described in Section 3 have their extreme values at the radial velocity zeros. Although Babcock (**4**) has pointed out that the values of H_p quoted have a probable error of about a kilogauss, it is nevertheless interesting to notice that the pulses in luminosity of the star take place during the apparently smaller variation in negative H_p. Thus the magnetic energy of the

star during the light maximum seems to be less than during the shallow light minimum when the apparently larger variation in positive H_p gives a larger magnetic energy of the star.

A theory of magnetic variability in stars has been given recently by Schwarzschild (**9**) in an investigation of the magneto-hydrodynamic oscillations which can take place in a non-rotating star composed of homogeneous incompressible matter with infinite conductivity. This idealized star is considered to have a permanent magnetic field on which is superimposed a small variable component h. The modes of oscillation investigated are those in which the vector potential of the solenoidal stream velocity is symmetrical and antisymmetrical with respect to the equatorial plane. These are called the even mode and the odd mode respectively. In the even mode the oscillatory component of the magnetic field has the character of a reversing quadrupole, and in the odd mode the field behaves like a reversing dipole. In both modes the zeros of stream velocity * occur at the phases of the amplitude maxima of the variable field. Although the theory in its present form is severely simplified it is interesting to note the agreement of the theory with the coincidence of the phases of zero velocity and the magnetic field maxima as found above for HD 125248.

We now examine some consequences of the light variability given in Section 6. We suppose that the luminosity of the star in the wave-length λ is given by $L_\lambda \propto R^2 B_\lambda(T)$, where R is the radius and $B_\lambda(T)$ the Planck Function. Since the photocell admits a fraction f of the intensity of radiation at the effective wave-length λ_e, the photometric measures in Table II give this reduced intensity $l = fL_{\lambda e}$ in terms of the maximum light as unit. If we neglect changes in the radius, thereby supposing that variations in the luminosity are brought about only by changes in T, we find

$$-\log_e l \sim \frac{hc}{k\lambda}\left(\frac{\Delta T}{T^2}\right).$$

With $T = 14{,}000$ deg. K. as the colour temperature for spectral type A0 the change ΔT in colour temperature required to account for the observed variations in luminosity at wave-length 4000 A. amounts to about 250 deg. K., the star being bluest at maximum light.

According to Struve and Swings (**3**) a change in colour temperature of about 2000 deg. K. has been found by Nikonov and Brodskaja (**10**) in α^2 C Ven, the star being bluest when the total light is a minimum. It should be pointed out that a colour temperature change of 2000 deg. K. would give a change of approximately 0·10 in the relative gradient. One should expect that a change in gradient by an amount as large as this would have shown up in the Greenwich gradients (**11**). But such an effect does not appear in the measurements, in fact α^2 C Ven is taken as one of the standard stars. More recently Stebbins and Whitford (**12**) have studied this star and did not find any large change in colour.

On the basis of a colour temperature change in HD 125248 of 250 deg. K., as suggested by the above calculation, a change in relative gradient of about 0·01 from maximum to minimum light might be expected.

6. This investigation benefited from the visit of Professor S. Chapman to the Commonwealth Observatory, and from some correspondence with

* Schwarzschild's stream velocities integrated over the visible hemisphere are taken to be the radial velocities measured.

Dr H. W. Babcock. The writer also acknowledges the benefit of conversations with his colleague Dr S. C. B. Gascoigne.

APPENDIX

Take the magnetic axis in the xz plane and inclined at an angle α to the axis of z, the line of sight. Let P be a point on the visible hemisphere with coordinates (x, y, z), and (X, Y, Z) with respect to the magnetic axes. The components of magnetic intensity at the point P are as follows:—

$$H_X = \frac{3M}{R^5} XZ, \qquad H_Y = \frac{3M}{R^5}, \qquad H_Z = \frac{M}{R^5}(2Z^2 - X^2 - Y^2),$$

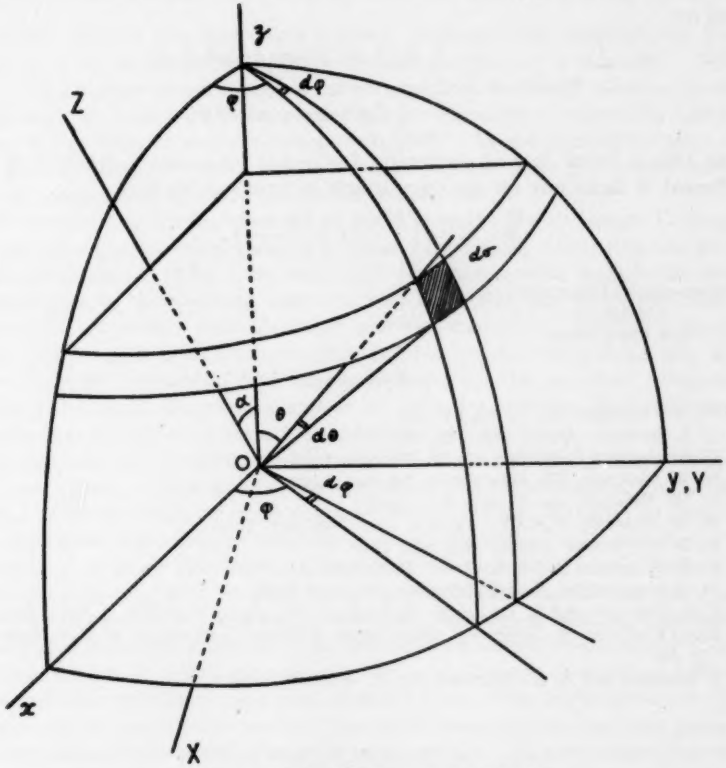


FIG. 4.—*Zeeman effect for an inclined dipole.*
Line of sight: oz.
Magnetic axis: OZ.

where M is the magnetic moment of the dipole at the centre of the star, and R is the radius. The component of magnetic force in the line of sight is

$$H_z = H_Z \cos\alpha - H_X \sin\alpha.$$

Introducing the transformation

$$X = x\cos\alpha - z\sin\alpha, \qquad Y = y, \qquad Z = z\cos\alpha + x\sin\alpha,$$

where $x = R\sin\theta\cos\phi$, $y = R\sin\theta\sin\phi$, $z = R\cos\theta$, we find

$$H_z = \frac{M}{R^3}(-\sin^2\theta\cos\alpha + 2\cos^2\theta\cos\alpha + 3\sin\theta\cos\theta\cos\phi\sin\alpha).$$

The displacement δ is proportional to the line of sight component of the magnetic intensity. To obtain the effect integrated over the visible hemisphere of a star which is darkened towards the limb, H_z has to be weighted by the line-of-sight flow of radiation $I_\lambda(\theta)\,d\sigma\cos\theta\,d\omega$ from the surface element $d\sigma$ at the point P. Since the area on the sphere between planes of co-latitude $\theta, \theta + d\theta$ is $2\pi R^2\sin\theta\,d\theta$, and the element of lune with angle $d\phi$ intercepts the fraction $d\phi/2\pi$ of this zonal area, we have $d\sigma = R^2\sin\theta\,d\theta\,d\phi$. Integrating over the visible hemisphere the displacement shown in the total light received from the star is given by

$$\delta(\alpha,\lambda) = a\frac{\int_0^{\pi/2}\int_0^{\pi} H_z(\theta,\phi)I_\lambda(\theta)\sin\theta\cos\theta\,d\theta\,d\phi}{\int_0^{\pi/2}\int_0^{\pi} I_\lambda(\theta)\sin\theta\cos\theta\,d\theta\,d\phi}.$$

If we take a linear law of darkening $I_\lambda(\theta) = I_0(1-u+u\cos\theta)$, where u is the coefficient of darkening for the wave-length in question, we find

$$\delta(\alpha,u) = a\left(\frac{M\cos\alpha}{10R^3}\right)\left(\frac{15+u}{3-u}\right).$$

Commonwealth Observatory,
Canberra:
1950 *February* 22.

References

(1) W. W. Morgan, *Ap. J.*, **74**, 24, 1931.
(2) A. J. Deutsch, *Ap. J.*, **105**, 283, 1947.
(3) O. Struve and P. Swings, *Ap. J.*, **98**, 361, 1943.
(4) H. W. Babcock, *The Observatory*, **69**, 191, 1949.
(5) H. W. Babcock, *Ap. J.*, **105**, 105, 1947.
(6) H. W. Babcock, *P.A.S.P.*, **59**, 260, 1947; **60**, 247, 1948.
(7) D. W. N. Stibbs, *Nature*, **165**, 195, 1950.
(8) A. R. Hogg and P. W. A. Bowe, *M.N.*, **110**, 373, 1950.
(9) M. Schwarzschild, *Ann. d'Astrophys.*, **12**, 148, 1948.
(10) V. B. Nikonov and A. Brodskaja, *Bull. Acad. Sci. Georgian S.S.R.*, **3**, 657, 1942.
(11) Royal Observatory, Greenwich, *Observations of Colour Temperatures of Stars*, 1926–32, p. 48.
(12) J. Stebbins and A. E. Whitford, *Ap. J.*, **102**, 273, 1945.

STELLAR PARALLAXES DETERMINED PHOTOGRAPHICALLY AT THE CAPE OBSERVATORY (SEVENTEENTH LIST)

(Communicated by H.M. Astronomer)

(Received 1950 April 17)

The principal list is in the same form as the previous lists beginning in *M.N.*, **89**, 497, 1929.

Nearly half of the stars were selected from the Cape Astrographic Zone $-40°$ to $-52°$ as stars with a proper motion exceeding $0''.2$ annually. Other stars with a similar proper motion were selected from the G.C. About a quarter of the stars are faint stars with a proper motion exceeding $0''.5$ given by Luyten. Many of the latter have a considerable parallax. As the interval between the first and last plates of a series is generally four years or more, the proper motion is more accurately determined than in the earliest lists.

The visual magnitudes were mostly taken from the Henry Draper Catalogue but the photographic magnitudes are new and depend on recent determinations made at the Cape. It has been found that the photographic magnitudes given in catalogues of the southern stars are very unreliable and in urgent need of observation by modern methods. For nearly all stars with a parallax exceeding $0''.100$ the error in determining a photographic absolute magnitude may arise more from the magnitude determination than from the parallax. Generally speaking the visual magnitudes appear to be more accurate than the photographic ones as judged by the colour index and spectral type.

In addition to the hundred stars for which a first determination of the parallax at the Cape is given, a list is given of stars for which a second determination has been made. These stars have generally been re-observed because of a large parallax, but for a few a re-determination has been made, as the earlier result disagreed with determinations made elsewhere. Completely new series of plates, considerably more extensive in general than the earlier series, have been used. In some cases a new set of comparison stars has been selected, although in most cases those previously used have been retained. The method of measurement used at the Cape is a differential one in which the positions of the parallax and comparison stars are measured relative to lines ruled on a reference plate. This method loses its value for stars of very large proper motion if the observations extend over many years, as the displacement may be rather large for measurement. In many cases it has been found convenient to make a new reference plate, although this means that the new series of measurements cannot be directly combined with the earlier ones for proper motion determination. Usually a few of the best of the earlier plates have been measured against the new reference plate. In all cases except No. 424 —40 Eridani A—the proper motion appears to be linear. In giving the results for these stars the original determination is given under (1), and later determinations under (2) or (3), an asterisk indicating that a new set of comparison stars has been used. In forming the means given for each star, consideration has been given to weight of the determination rather than the individual probable error.

Notes

1501. The spectral type is from a Bloemfontein plate. Cillié writes: Relative intensities of the G band and *Ca* I 4226 make it K2. The *Mn* I (4030–4034) and nearby *Fe* I and *K* I lines appear more intense than they should be.
1504. Preceding, north and fainter star of a pair. The brighter star is No. 602: P.G.C. 285, parallax $+0''.078 \pm 0''.011$.
1516. The position does not agree very well with that of CoD $-51°$ 1204. The CPD number is 569.
1529. LPM 269 is a white dwarf.
1539. This is a double star, separation about $2''$; the measurement of the plates was difficult and was done by two measurers independently.
1550. A difficult double star.
1559. There is a considerable difference between the relative and absolute proper motion.
1577. This star was put on the list as a double star with a large dynamical parallax. Apparently it is an optical double.

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0)	Proper Motion R.A.	Proper Motion Dec.	P.E. Unit Weight	Parallax	P.E.
				h m s					
1501	LPM 7	9·8	K2	0 06 31	+ 710	− 90	±26	+ 39	± 7
	−39° 31	10·96		−39° 47′·3	+ 730				± 3
1502	FS 211	...	...	0 42 10	+ 396	+116	±23	− 23	± 7
	−41° 200	11·92		−41° 21′·2	+ 375				± 3
1503	PGC 255	3·60	Ko	1 03 34	+ 213	−132	±32	+ 63	± 9
	GC 1384	4·55		−10° 42′·7	+ 214				± 4
1504	GC 1490	8·7	G7	1 09 20	+ 131	+284	±27	+ 64	± 7
	−8° 215	8·47		− 8° 26′·9	+ 86				± 4
1505	FS 363	...	...	1 13 39	+ 220	−111	±32	+ 43	± 8
	CAZ −49° 25	12·86		−48° 40′·8	+ 190				± 4
1506	LPM 96	...	Mo	2 00 10	+1290	−180	±25	+127	± 6
	−18° 359	11·58		−18° 05′·4	+1296				± 3
1507	Ci 18 296	8·9	Go	2 13 26	+ 305	− 49	±32	+ 17	± 8
	−7° 397	8·85		− 7° 04′·0	+ 281				± 4
1508	Ci 18 301	8·2	Go	2 14 59	+ 214	+ 70	±28	+ 23	± 8
	−1° 317	8·81		− 0° 55′·2	+ 172				± 4
1509	GC 3106	6·78	Go	2 30 19	− 150	−427	±20	+ 43	± 6
	79 Ceti	7·39		− 3° 59′·2	− 156				± 3
1510	PGC 612	5·73	F5	2 36 07	+ 214	−130	±30	+ 39	± 7
	GC 3235	6·15		− 1° 07′·3	+ 214				± 3
1511	PGC 665	4·05	Ko	2 51 33	+ 78	−213	±28	+ 37	± 7
	GC 3539	4·97		− 9° 17′·8	+ 59				± 3
1512	PGC 750	5·05	Ko	3 13 57	+ 20	+ 15	±21	+ 20	± 6
	GC 3955	5·68		−22° 52′·6	+ 13				± 4
1513	GC 4975	8·39	Ko	4 03 15	− 69	−282	±33	+ 26	± 8
	−27° 1560	9·17		−27° 41′·1	− 97				± 6
1514	FS 1213	10·0	K2	4 [illegible]5 28	+ 24	+139	±20	+ 10	± 6
	−43° 1321	10·91		−43° 37′·7	+ 13				± 5
1515	GC 5542	7·19	K5	4 28 07	+ 173	−108	±25	+ 22	± 6
	−29° 1770	8·37		−29° 01′·0	+ 153				± 4

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0)	Proper Motion R.A.	Proper Motion Dec.	P.E. Unit Weight	Parallax	P.E.
				h m s					
1516	FS 1404	...	...	4 36 14	+ 168	+ 81	±22	+ 11	± 6
	−51° 1204	11·30		−51° 54′·8	+ 169				± 4
1517	GC 5704	7·65	Go	4 36 43	− 9	−283	±34	+ 16	± 8
	−28° 1674	8·16		−28° 23′·3	− 29				± 5
1518	GC 5843	5·63	Go	4 43 07	+ 133	+174	±36	+ 96	± 9
	−17° 954	5·93		−17° 07′·1	+ 128				± 5
1519	LPM 229	...	G5	6 04 46	− 180	−720	±26	+ 44	± 7
	−59° 584	9·10		−59° 30′·2	− 227				± 4
1520	PGC 1589	5·14	Ko	6 13 13	+ 122	−218	±29	+105	± 8
	GC 7962	5·64		−74° 43′·1	+ 120				± 5
1521	LPM 241	...	m	6 32 23	− 410	+770	±22	+ 69	± 7
		12·91		−58° 27′·8	− 409				± 3
1522	GC 8954	5·83	F2	6 45 44	+ 30	−185	±40	+ 29	±10
	−0° 1462	5·97		− 0° 25′·1	+ 33				± 7
1523	PGC 1769	3·30	A5	6 47 10	− 74	+262	±23	+ 48	± 6
	GC 8941	3·33		−61° 50′·0	− 73				± 4
1524	GC 9106	8·37	G5	6 52 51	− 44	−193	±19	+ 67	± 5
	−55° 1095	8·75		−55° 07′·5	− 42				± 3
1525	LPM 251	...	f-g	7 04 43	− 80	+670	±20	+ 30	± 5
	−57° 1139	9·82		−57° 19′·2	− 84				± 4
1526	FS 2914	...	Go	7 11 44	− 29	+175	±14	+ 14	± 4
	−43° 3055	10·27		−43° 29′·9	− 29				± 2
1527	GC 9763	7·14	F5	7 15 58	− 235	+118	±37	+ 7	± 9
	−45° 3088	7·45		−45° 07′·0	− 233				± 3
1528	FS 2967	...	...	7 16 09	+ 202	−103	±21	+ 29	± 5
	−40° 3074	11·54		−41° 02′·3	+ 194				± 2
1529	LPM 269	...	b	7 35 44	+1150	−590	±22	+151	± 6
		13·08		−17° 09′·9	+1138				± 3
1530	GC 10418	7·9	Go	7 39 23	− 9	+189	±38	+ 30	±10
	−4° 2069	8·15		− 4° 49′·2	− 39				± 5
1531	Ross 884	...	...	7 48 31	+ 150	−410	±34	+ 28	±10
		12·98		− 6° 28′·1	+ 100				± 4
1532	PGC 2089	4·53	B3	7 49 06	− 11	− 8	±23	+ 9	± 6
	GC 10661	4·10		−38° 36′·2	− 15				± 3
1533	FS 3513	...	K	7 50 17	− 25	+169	±37	+ 14	±10
	−49° 3139	12·50		−49° 55′·4	− 18				± 4
1534	ZC 5038	...	Ko	7 52 00	− 111	+260	±29	+ 25	± 7
	−49° 3174	10·11		−49° 34′·7	− 103				± 3
1535	PGC 2163	5·64	Go	8 06 02	− 243	+ 54	±32	+ 59	± 7
	GC 11118	5·86		−13° 30′·3	− 255				± 4
1536	LPM 283	...	k-m	8 08 30	− 220	−470	±29	+ 76	± 7
	−13° 2439	10·77		−13° 36′·6	− 219				± 5

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0) h m s	Proper Motion R.A.	Proper Motion Dec.	P.E. Unit Weight	Parallax	P.E.
1537	LPM 292	...	k	8 24 34	+ 410	−870	±30	+ 51	± 7
		12·68		− 1° 23′·4	+ 408				± 5
1538	LPM 298	8·95	G5	8 45 24	− 110	−520	±30	+ 17	± 7
	−4° 2468	9·74		− 5° 08′·9	− 153				± 4
1539	LPM 299	...	M3	8 49 18	+ 360	−500	±21	+ 79	± 6
		13·35		−12° 43′·9	+ 359				± 3
1540*	LPM 300	...	g	8 54 04	+ 230	−490	±38	+ 10	±10
		12·34		− 5° 59′·6	+ 221				± 7
1541	GC 12632	7·31	Go	9 04 02	− 375	−128	±34	+ 36	± 8
	−10° 2754	7·60		−10° 21′·0	− 339				± 6
1542	LPM 316	11·6	M4	9 23 58	− 120	−640	±31	+ 78	±10
	Ross 439	13·52		− 6° 54′·7	− 173				± 4
1543	ZC 7700	8·6	F8	9 39 10	− 130	+118	±27	+ 28	± 7
	−42° 5552	8·69		−42° 49′·1	− 110				± 2
1544	GC 13670	7·68	Go	9 51 05	+ 169	−116	±20	+ 35	± 5
	−20° 3049	8·14		−20° 16′·2	+ 196				± 3
1545	LPM 337	...	m	10 07 25	− 520	0	±28	+ 60	± 8
	−17° 3088	11·33		−18° 07′·5	− 528				± 4
1546	GC 14077	6·72	Ao	10 10 37	− 113	+ 30	±32	− 2	± 8
	−41° 5716	6·54		−41° 13′·8	− 87				± 4
1547†	FS 7487	...	...	10 10 43	− 168	+ 51	±30	+ 11	± 8
	−40° 4306	11·91		−41° 05′·8	− 132				± 3
1548	LPM 347	...	K2	10 20 18	− 710	+100	±33	+ 63	± 8
	−9° 3070	11·52		− 9° 43′·6	− 823				± 5
1549	LPM 349	...	k	10 21 39	− 10	−590	±41	+ 33	± 9
	−5° 3063	10·62		− 6° 00′·1	+ 12				± 5
1550	FS 7920	...	G5	10 27 50	− 94	−133	±37	+ 16	± 9
	−41° 5939	10·68		−41° 37′·8	− 52				± 4
1551	LPM 353	...	m	10 34 42	− 670	−110	±25	+100	± 6
		12·69		− 6° 24′·1	− 695				± 4
1552	LPM 357	...	m	10 40 56	−1860	−610	±26	+ 91	± 7
	−18° 3019	12·40		−18° 34′·8	−1861				± 5
1553	Ross 898	...	...	10 49 58	− 310	+ 60	±30	+ 34	± 7
	−0° 2387	11·05		− 0° 16′·3	− 221				± 3
1554	GC 15212	6·94	A5	10 59 19	− 213	+ 29	±35	+ 15	± 9
	−50° 5641	6·93		−50° 48′·9	− 204				± 4
1555	FS 8569	...	k	10 59 35	− 238	− 3	±39	+ 28	±10
	−45° 6641	12·02		−45° 55′·6	− 243				± 4
1556	FS 8645	...	...	11 04 28	− 183	+146	±25	+ 27	± 6
	−41° 6360	10·92		−41° 18′·8	− 180				± 2

* 1540 is not −5° 2678 but about 1½′ np to that star.
† 1547 CPD Number. Not in CoD.

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0) h m s	Proper Motion R.A.	Proper Motion Dec.	P.E. Unit Weight	Parallax	P.E.
1557	LPM 370	...	K5	11 06 14	− 870	+660	±25	+ 48	± 7
	−10° 3216	10·26		−10° 25′·4	− 951				± 4
1558	LPM 388	...	m	11 27 01	− 670	+240	±33	+ 82	±10
	CAZ −41° 627	12·89		−40° 30′·0	− 697				± 6
1559	PGC 3058	4·47	Ko	11 31 50	+ 3	+ 38	±28	+ 7	± 7
	υ Leonis	5·26		− 0° 16′·3	+ 66				± 3
1560*	CoD −55° 4223	...	...	11 40 37	− 331	0	±36	+ 24	±10
		11·82		−55° 08′·6	− 309				± 2
1561	Ross 920	...	m	11 56 02	− 460	+200	±30	+ 44	± 8
	LPM 410	12·02		− 1° 10′·7	− 460				± 4
1562p	GC 16729	8·3	...	12 10 00	− 228	− 76	±36	+ 56	±10
	− 6° 3532p	8·86		− 6° 42′·0	− 216				± 5
1562f	GC 16730	8·0	G5	12 10 00	− 237	− 61	±33	+ 22	± 9
	− 6° 3532f	8·62		− 6° 42′·0	− 227				± 5
	Mean							+ 39	± 7
					− 222				± 4
1563	Ross 935	...	...	12 11 56	− 340	−160	±32	− 7	± 8
	−3° 3258	10·86		− 3° 41′·5	− 220				± 3
1564	Ross 938	...	...	12 23 40	− 280	+350	±26	+ 21	± 8
	−6° 3580	11·84		− 6° 58′·0	− 259				± 3
1565	ZC 10450	9·8	K5	12 25 59	− 421	−203	±27	+ 22	± 7
	−42° 7705	10·87		−42° 49′·6	− 382				± 4
1566	LPM 447	...	k	12 39 30	− 450	−180	±28	+ 1	± 7
	Ross 704	12·93		−14° 49′·0	− 441				± 2
1567	Ross 970	...	...	12 49 48	− 220	−200	±31	+ 42	± 8
	−5° 3596	11·43		− 5° 47′·5	− 286				± 4
1568	LPM 456	...	m	12 55 12	− 350	−410	±35	+ 27	± 8
		12·38		−62° 38′·6	− 376				± 5
1569	GC 17784	9·02	G5	13 01 51	− 228	− 54	±30	+ 26	± 7
	−49° 7651	9·36		−49° 48′·9	− 232				± 3
1570	FS 10788	...	G5	13 07 37	− 155	−144	±28	+ 14	± 7
	−40° 7755	11·99		−41° 04′·6	− 147				± 3
1571	PGC 3452	2·91	A2	13 14 58	− 339	− 92	±37	+ 45	± 8
	ι Centauri	2·62		−36° 11′·1	− 329				± 4
1572	GC 18284	7·31	G5	13 26 36	− 831	+254	±35	+ 50	± 9
	− 1° 2832	7·94		− 1° 48′·6	− 847				± 3
1573	ZC 11590	8·4	Ko	13 37 28	− 208	−101	±31	− 9	± 7
	−50° 7991	8·90		−51° 07′·2	− 190				± 5
1574	LPM 503	...	k	13 55 14	− 670	+110	±26	+ 21	± 6
	−31° 10833	12·36		−31° 19′·0	− 696				± 4

* 1560 Stars No. 453, 454 in this neighbourhood are brighter stars mistaken for this. They are CoD −55° 4219, 4220, or CPD −55° 4608, 4610.

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0) h m s	Proper Motion R.A.	Proper Motion Dec.	P.E. Unit Weight	Parallax	P.E.
1575	LPM 504	...	K5	13 55 58	− 710	+550	±32	+ 74	± 8
	−1° 2892	11·04		− 2° 11′·2	− 828				± 6
1576	FS 11966	...	...	14 10 12	− 215	−132	±25	+ 9	± 6
	−47° 9018	12·02		−48° 06′·4	− 197				± 3
1577	ADS 9476	9·2	K2	14 55 57	...	...	±39	+ 12	± 9
	−3° 3707	10·01		− 3° 42′·1	+ 3				± 4
1578	FS 13038	...	F8	15 15 35	− 145	−152	±28	− 2	± 7
	−50° 9325	11·03		−50° 57′·2	− 138				± 4
1579	ZC 13716	...	G5	15 32 45	− 131	−127	±21	+ 36	± 5
	−51° 9352	10·38		−51° 39′·6	− 126				± 2
1580	GC 21104	8·90	A0	15 35 59	− 165	− 84	±26	+ 8	± 6
	−44° 10333	8·84		−44° 37′·3	− 180				± 3
1581	GC 21518	7·17	F5	15 55 17	− 20	−165	±34	+ 24	± 7
	−42° 10980	7·35		−42° 24′·9	+ 6				± 4
1582	FS 14300	...	...	15 55 29	− 122	−192	±17	+ 8	± 5
	CAZ −50° 112	12·44		−49° 49′·9	− 118				± 2
1583	FS 14594	...	K0	16 11 14	+ 27	−215	±19	+ 24	± 4
	−46° 10628	10·71		−46° 32′·0	+ 22				± 2
1584	ZC 14907	8·6	G5	16 26 18	0	−255	±36	+ 14	± 7
	−45° 10738	8·95		−45° 34′·1	+ 13				± 3
1585	FS 15119	...	K0	16 27 39	− 228	− 85	±35	+ 27	± 9
	−45° 10752	10·45		−45° 13′·0	− 200				± 5
1586	FS 15913	...	G0	16 55 58	− 52	−170	±23	− 8	± 5
	−44° 11376	10·80		−44° 34′·7	− 68				± 3
1587	FS 17568	...	K2	17 50 04	+ 52	−190	±22	+ 12	± 6
	−46° 11973	11·38		−46° 31′·6	+ 41				± 3
1588	ZC 17382	8·1	G0	18 19 14	+ 143	−146	±31	+ 25	± 7
	−51° 11566	8·45		−51° 56′·3	+ 126				± 3
1589	FS 18877	...	G5	19 44 51	+ 212	−211	±26	0	± 6
	−41° 13726	11·42		−41° 17′·7	+ 185				± 4
1590	FS 19269	...	...	20 39 20	+ 7	−502	±39	+ 28	±10
	−47° 13548	13·02		−47° 30′·5	+ 17				± 5
1591	FS 19299	...	...	20 46 29	− 31	−253	±29	+ 13	± 7
	CAZ −51° 330	12·50		−51° 06′·4	− 29				± 3
1592	GC 29345	6·51	G5	20 55 36	+ 74	−108	±38	+ 14	± 9
	−43° 14325	7·19		−43° 23′·5	+ 68				± 4
1593	GC 29346	6·94	G5	20 55 41	+ 67	−104	±31	− 4	± 8
	−43° 14327	7·79		−43° 23′·2	+ 55				± 4
1594	ZC 19282	9·4	G0	21 09 31	+ 74	−190	±28	+ 7	± 7
	−49° 13403	10·11		−49° 51′·0	+ 96				± 4
1595	PGC 5765	5·82	F0	22 17 00	+ 229	− 49	±20	+ 25	± 5
	GC 31279	5·90		−46° 25′·9	+ 211				± 3

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0)	Proper Motion R.A.	Proper Motion Dec.	P.E. Unit Weight	Parallax	P.E.
				h m s					
1596	FS 20044	...	K2	22 37 19	+ 236	− 94	±33	+ 11	± 9
	CAZ −48° 53	12·26		−47° 20′·3	+ 208				± 6
1597	GC 32094	7·74	G5	22 56 41	+ 197	−216	±25	+ 9	± 6
	−50° 13857	8·51		−50° 13′·7	+ 182				± 3
1598	FS 20279	...	K5	23 12 04	+ 294	− 49	±18	+ 58	± 5
	−42° 16263	11·61		−42° 44′·2	+ 296				± 3
1599	ZC 20498	9·16	Ko	23 22 48	− 4	−305	±17	+ 9	± 4
	−45° 15036	9·89		−45° 00′·2	− 35				± 2
1600	FS 20521	...	...	23 52 00	+ 231	− 77	±31	+ 27	± 7
	CAZ −48° 131	12·00		−47° 11′·6	+ 209				± 3

Second Determinations

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0)	P.M. in R.A.	Relative	Parallax	P.E.
				h m s				
1	GC 338	7·97	Go	0 12 14	+349± 8	− 3	± 8	(1)
	−44° 52	8·62		−44° 24′·5	+386± 2	+ 28	± 9	(2)
					+384± 2	+ 18	± 7	Mean
215	GC 354	6·62	Go	0 12 51	+437± 9	+ 1	± 9	(1)
	−14° 42	6·96		−14° 00′·7	+411± 2	+ 14	±10	(2)
					+412± 2	+ 8	± 7	Mean
114	−24° 889	9·5	G5	2 00 58	+368± 7	+ 16	± 7	(1)
		9·69		−24° 51′·4	+398± 2	− 5	±10	(2)
					+396± 2	+ 6	± 7	Mean
115	−24° 891	9·2	G5	2 01 02	+404± 9	− 2	± 9	(1)
		9·35		−24° 51′·2	+400± 2	− 3	±10	(2)
					+400± 2	− 2	± 9	Mean
224	GC 2794	6·38	Ko	2 14 30	−225± 8	+ 61	± 8	(1)
	−26° 828	6·88		−26° 25′·1	−247± 1	+ 74	± 5	(2)
					−245± 1	+ 64	± 7	(3)
					−246± 1	+ 69	± 4	Mean
225	GC 3097	8·83	K5	2 30 34	− 4±10	+ 77	±10	(1)
	−44° 775	10·12		−44° 13′·6	+ 37± 1	+ 88	± 8	(2)*
					+ 37± 1	+ 84	± 7	Mean
121	GC 3279	5·42	F5	2 39 09	+329± 9	+ 73	± 8	(1)
	−51° 641	5·47		−51° 13′·9	+331± 3	+ 42	± 6	(2)
					+336± 1	+ 53	± 5	Mean
123	GC 3449	6·14	G5	2 47 43	+402± 8	+131	± 8	(1)
	−13° 544	6·76		−13° 10′·5	+381± 1	+109	± 7	(2)*
					+381± 1	+118	± 5	Mean
126	GC 3641	5·90	G5	2 57 19	+257± 8	+ 15	± 9	(1)
	−28° 987	6·52		−28° 28′·5	+292± 2	+ 35	± 9	(2)
					+290± 1	+ 25	± 7	Mean
850	FS 821	...	...	2 59 03	+186± 5	+ 40	± 5	(1)
	CAZ −46° 110	12·23		−45° 02′·1	+167± 5	+ 11	± 8	(2)
					+173± 2	+ 26	± 5	Mean

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0) h m s	P.M. in R.A.	Relative	Parallax	P.E.
127	GC 3966	5·48	Go	3 15 36	+1327± 7	+107	± 7	(1)
	−63° 217	5·99		−62° 57′·5	+1327± 1	+104	± 6	(2)
					+1327± 1	+105	± 5	Mean
128	GC 3975	5·16	Go	3 16 02	+1304± 6	+ 98	± 7	(1)
	−62° 265	5·68		−62° 53′·3	+1326± 1	+ 63	± 8	(2)
					+1326± 1	+ 82	± 6	Mean
518	GC 4778	3·18	K5	3 53 22	+ 62±11	+ 14	±10	(1)
856	−13° 781	4·62		−13° 47′·6	+ 26± 4	+ 8	± 6	(2)*
					+ 48± 2	+ 8	± 8	(3)
					+ 52± 2	+ 9	± 5	Mean
424	GC 5138	4·48	G5	4 10 40	−2274± 7	+191	± 7	(1)
	−7° 780	5·11		− 7° 48′·5	−2286± 6	+216	± 7	(2)*
					−2236± 2	+204	± 5	Mean
329	GC 5140	9·7	A	4 10 46	−2177± 8	+183	±10	(1)
	−7° 781	9·18		− 7° 48′·8	−2234± 7	+182	± 8	(2)*
					−2234± 2	+182	± 6	Mean
672	GC 5428	6·18	F5	4 24 10	+ 64±11	+ 22	±10	(1)
	−47° 1383	6·39		−47° 09′·8	+ 26± 2	+ 33	± 7	(2)*
					+ 27± 2	+ 29	± 6	Mean
147	GC 5824	7·77	G5	4 43 7	− 485± 8	+ 57	± 8	(1)
	−50° 1492	8·34		−50° 14′·7	− 468± 1	+ 56	± 5	(2)
					− 474± 2	+ 42	± 8	(3)
					− 468± 1	+ 52	± 5	Mean
150	GC 6120	6·50	Ko	4 55 51	+ 555± 8	+ 88	± 9	(1)
	−5° 1123	7·12		− 5° 52′·3	+ 546± 4	+139	± 9	(2)*
					+ 548± 4	+114	± 6	Mean
339	GC 7196	6·41	G5	5 40 16	− 318±10	+134	±10	(1)
	−22° 1210	6·93		−22° 27′·3	− 308± 2	+109	± 9	(2)
					− 309± 2	+120	± 7	Mean
428	GC 7197	3·80	F8	5 40 18	− 299± 8	+134	± 8	(1)
	−22° 1211	3·94		−22° 28′·9	− 294± 4	+122	±10	(2)
					− 295± 4	+128	± 6	Mean
176	GC 11115	4·80	F5	8 07 21	− 172± 7	+ 50	± 6	(1)
	−60° 1074	5·00		−60° 59′·9	− 172± 8	+ 48	± 8	(1)*
					− 156± 1	+ 46	± 7	(2)*
					− 156± 1	+ 48	± 5	Mean
367	FS 6566	...	Ko	9 40 44	− 455± 7	+101	± 9	(1)
	LPM 328	11·70		−45° 18′·0	− 452± 3	+123	± 9	(2)*
					− 453± 3	+112	± 6	Mean
705	FS 20559	11·6	M4+	9 54 34	+ 523±10	+ 16	±12	(1)
	−45° 5627	12·58		−45° 56′·0	+ 506± 6	+ 78	±10	(2)
					+ 496± 2	+ 63	± 7	Mean
625	GC 15693	5·34	F5	11 21 23	− 289± 9	+ 41	± 8	(1)
	−63° 1893	5·69		−63° 25′·2	− 318± 2	+ 18	±10	(2)
					− 316± 2	+ 26	± 7	Mean
730	PGC 3448	4·80	G5	13 13 10	−1056± 9	+ 88	± 8	(1)
	61 Virginis	5·39		−17° 45′·3	−1049± 3	+ 85	±10	(2)
					−1050± 3	± 87	± 6	Mean

No.	Name and D.M. No.	Mag. Vis. Phot.	Type	R.A., Dec. (1900.0)	P.M. in R.A.	Relative	Parallax	P.E.
				h m s				
58	GC 20111	8·87	M2	14 51 36	+ 988± 7	+148	± 6	(1)
	−20° 4123	9·56		−20° 57′·7	+1000± 6	+154	± 7	(2)*
					+ 994± 4	+151	± 4	Mean
185	GC 20113	5·76	K5	14 51 37	+1025± 7	+172	± 6	(1)
	−20° 4125	6·65		−20° 57′·9	+1030± 5	+158	± 6	(2)*
					+1028± 4	+165	± 4	Mean
562	FS 13312	8·7	Mo	15 25 44	−1163±10	+161	± 8	(1)
	−40° 9712	10·84		−40° 54′·4	−1197± 2	+148	± 8	(2)*
					−1195± 2	+154	± 6	Mean
268	GC 24802	8·26	Ma	18 05 06	+ 119± 6	+ 76	± 5	(1)
	ZC 17132	9·56		−43° 27′·3	+ 130± 1	+ 44	± 5	(2)*
					+ 130± 1	+ 55	± 4	Mean
69	GC 25424	7·44	Go	18 31 01	+ 148±10	+ 51	± 9	(1)
	−25° 13291	7·93		−25° 44′·8	+ 157± 1	+ 36	± 5	(2)
					+ 157± 1	+ 41	± 4	Mean
597	GC 32863	6·7	Ko	23 33 42	+ 94± 5	+133	± 7	(1)
	−73° 2299	8·01		−73° 15′·3	+ 139± 2	+103	± 9	(2)*
					+ 133± 2	+118	± 6	Mean

NOTES

224. ($2^h\ 14^m$) (2) includes (1).

176. ($8^h\ 07^m$) (1)* is deduced from the same measures as (1) but with two of the four reference stars rejected as in (2).

ERRATA

M.N., **109,** No. 6, Harold Jeffreys' paper:

P. 678, equations (27) and (28), *for* L *read* L/n.

P. 687, line 1 of footnote, *for* Takenti *read* Takeuti.

CONTENTS

PAGE

Meeting of 1950 May 12 :

Fellows elected 263
Presents announced... 263
George Darwin Lecture announced 263

Additional Meeting of 1950 September 22 :

Tribute to the late Professor E. A. Milne 263
Presents announced 265
Other Dublin meetings 265

E. A. Milne, Gravitation and magnetism 266

C. M. Bondi, Models for red giant stars. Paper I. General discussion and application to homogeneous models 275

C. M. Bondi and H. Bondi, Models for red giant stars. Paper II. Models with a chemical inhomogeneity and opacity due to photoelectric effect 287

G. L. Camm, Self-gravitating star systems 305

W. H. Ramsey, On the instability of small planetary cores.—I 325

M. J. Lighthill, On the instability of small planetary cores.—II 339

A. D. Thackeray, Observations of the brightest regions of three diffuse nebulae ... 343

R. v. d. R. Woolley and C. W. Allen, Ultra-violet emission from the chromosphere 358

A. R. Hogg and P. W. A. Bowe, Photoelectric observations of S Antliae 373

M. Ryle and A. Hewish, The effects of the terrestrial ionosphere on the radio waves from discrete sources in the galaxy 381

D. W. N. Stibbs, A study of the spectrum and magnetic variable star HD 125248... 395

Royal Observatory, Cape of Good Hope, Stellar parallaxes determined photographically at the Cape Observatory (seventeenth list) 405

Errata 413

Printed in London by Taylor & Francis, Ltd., 1951 February